

OM nucleic - nucleic search, using sw model
Run on: June 7, 2004, 08:48:09 ; Search time 6508 Seconds
(without alignments)
10982.278 Million cell updates/sec

Title: US-09-763-822A-13
Perfect score: 1649
Sequence: 1 ccatggcacttacagaagaa.....ctatgagatatatgaggatccc 1649

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.mu.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg.hum.*

31: em.htg.inv.*

32: em.htg.other.*

33: em.htg.mus.*

34: em.htg.pln.*

35: em.htg.rod.*

36: em.htg.mam.*

37: em.htg.vrt.*

38: em.sy.*

39: em.htg.hum.*

40: em.htg.mus.*

41: em.htg.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1640.6	99.5	2076	8	AY006482	AY006482 Artemisia
2	1637.4	99.3	2076	6	BD227440	BD227440 Transgeni
3	1635.8	99.2	2080	8	AF138959	AF138959 Artemisia
4	1634.2	99.1	2106	8	AAN251751	AF251751 Artemisia
5	1626.2	98.6	2067	8	AF327526	AF327526 Artemisia
6	857	52.0	1902	8	AF472361	AF472361 Artemisia
7	836	50.7	2026	8	AAN271793	AJ271793 Artemisia
8	776.2	47.1	1976	8	SCA304452	AJ304452 Solidago
9	723.4	43.9	1946	8	AATERPENE	AJ001539 Artemisia
10	723.4	43.9	1985	8	AF157059	AF157059 Artemisia
11	723.2	43.9	4392	8	AF327527	AF327527 Artemisia
12	721.8	43.8	1893	8	AAN249561	AJ249561 Artemisia
13	588	35.7	1886	8	AF304444	AF304444 Artemisia
14	559	33.9	1843	8	AAN271792	AJ271792 Artemisia
15	510.6	31.0	538	6	BD227439	BD227439 Transgeni
16	424.6	25.7	1890	8	GACAD1C2	Y16432 Gossypium a
17	424.6	25.7	1893	8	GAU23205	U23205 Gossypium a
18	418.2	25.4	1820	8	AF270425	AF270425 Gossypium
19	418.2	25.4	1892	8	GAU23206	U23206 Gossypium a
20	416.6	25.3	1851	8	GHU88318	U88318 Gossypium h
21	399.6	24.2	1883	8	GAU27535	U27535 Gossypium a
22	399.6	24.2	1935	8	GACADSYNH	X96429 Gossypium a
23	396.2	24.0	1973	6	AX529025	AX529025 Sequence
24	390.2	23.7	1900	8	AF288465	AF288465 Citrus ju
25	388.4	23.6	2030	8	AF279455	AF279455 Lycopersi
26	386.4	23.4	1647	8	AF441124	AF441124 Citrus si
27	383.2	23.2	1877	8	AF411120	AF411120 Citrus x
28	369.8	22.4	2605	6	AX529022	AX529022 Sequence
29	364	22.1	1879	6	AR183920	AR183920 Sequence
30	364	22.1	1879	8	AF035630	AF035630 Lycopersi
31	364	22.1	2003	8	AF279454	AF279454 Lycopersi
32	364	22.1	2018	8	AF279453	AF279453 Lycopersi
33	363.6	22.0	1672	6	AX529019	AX529019 Sequence
34	361.2	21.9	2000	8	AF279456	AF279456 Lycopersi
35	357.6	21.7	2024	6	BD227674	BD227674 Synthesas
36	357.6	21.7	2024	6	AR183921	AR183921 Sequence
37	357.6	21.7	2024	6	AR240709	AR240709 Sequence
38	357.6	21.7	2024	6	AR266984	AR266984 Sequence
39	357.6	21.7	2024	6	AR316333	AR316333 Sequence
40	357.6	21.7	2024	6	AR338476	AR338476 Sequence
41	357.6	21.7	2024	6	AR429882	AR429882 Sequence
42	357.6	21.7	2024	8	AF035631	AF035631 Lycopersi
43	340	20.6	1867	8	AF374462	AF374462 Citrus ju
44	330.4	20.0	1944	6	BD227667	BD227667 Synthesas
45	330.4	20.0	1944	6	AR240702	AR240702 Sequence

ALIGNMENTS

RESULT 1
AY006482
LOCUS AY006482 2076 bp mRNA linear PLN 03-JUL-2001
DEFINITION Artemisia annua amorpho-4,11-diene synthase mRNA, complete cds.
ACCESSION AY006482
VERSION AY006482.1 GI:9837803
KEYWORDS Artemisia annua (sweet wormwood)
ORGANISM Artemisia annua
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Asteroideae;
Anthemideae; Artemisia.
REFERENCE 1 (bases 1 to 2076)
AUTHORS Wallaart,T.E., Bouwmeester,H.J., Hille,J., Poppinga,L. and

Maïjers, N.C.
Amorpha-4,11-diene synthase: cloning and functional expression of a
key enzyme in the biosynthetic pathway of the novel antimalarial
drug artemisinin
Planta 212 (3), 460-465 (2001)
21183247
11289612
2 (bases 1 to 2076)
Wallaart, T.E. and Bouwmeester, H.J.
DNA encoding amorpha-4,11-diene synthase
Patent: European Patent Office (98202854.0) - EPO Bulletin 2000/09 ;
GenoClipp Biotechnology B.V. ;
The Netherlands ;
3 (bases 1 to 2076)
Wallaart, T.E., Bouwmeester, H.J., Hille, J., Poppinga, L. and
Maïjers, N.C.A.
Direct Submission
Submitted (09-AUG-2000) GenoClipp Biotechnology B.V., Meditech
Center, L.J. Zielstraweg 1, Groningen 9713 GX, The Netherlands
Location/Qualifiers
1. .2076
/organism="Artemisia annua"
/mol_type="mRNA"
/db_xref="taxon:35608"
6. .1646
/codon_start=1
/product="amorpha-4,11-diene synthase"
/protein_id="AAF98444.1"
/db_xref="GI:3837804"
/translation="MSLITEEPPIPIANFPSPINGDQFLIYQKQVEQVQVINDLKK
EVRLLEALDIPMKHANLLKLDIEIQRLGIPYHFERIDHALQCIETYGDNMNGDR
SDALFRLRKGYYVTCDFVFNKQKQAFKQSLANDVEGLLELYEATSMRVPGBIIL
EDALGFTSRSLINIKDQAFSTNPALFTIQRALQPLWKRLPRIEAAQYIIPYQOQDS
HNKTLLEKLEFNLKSHLESHVCKWKAFDIKNAQPLRDRIVECVFWGLSGG
YEPQYSARVPETKAVAVITLIDDTYDAYGYEELKIPTFAVERBSIICLDTLPEYMK
PIYKLFMDTYTEMEFLAKGRDLPNGCKEFPVEFVNLNMEVAKWANEHPIPTTEHK
DPVITLGGANLLTYTCLGMSDIFTKESVEWASAPPLFRYGLIGRLRLNDLMTKHA
EQERKHSLSLSELYNNENYEAOTLIYKEVEDVWPKDINREYLTITKNIIPRLLMAY
IYLQFLEVOYAGKDNFTRMGDYKHLIKSLLVPMST"

ORIGIN

Query Match 99.5%; Score 1640.6; DB 8; Length 2076;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1643; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 CATGGCACTTACAGAGAAAACCTATTTCGCCCACTTCCCAACTTTCCTCCAAAGCATTTG 61
DB 5 CATGTCACTTACAGAGAAAACCTATTTCGCCCACTTCCCAACTTTCCTCCAAAGCATTTG 64
QY 62 GGGAGATCAGTTTCTCATCTATCAAAACCAAGTAGAGCAAGGGTGGAAACAGATAGTGAA 121
DB 65 GGGAGATCAGTTTCTCATCTATCAAAACCAAGTAGAGCAAGGGTGGAAACAGATAGTGAA 124
QY 122 TGATTTAAAAAGAGTGGCGCACTACTTAAAGAGCTTTGGATATTCCTATGAACA 181
DB 125 TGATTTAAAAAGAGTGGCGCACTACTTAAAGAGCTTTGGATATTCCTATGAACA 184
QY 182 TGCCAAATTTGTTGAAGCTGATTGATGAATTCACGCCCTTGGAAATPACCGTATCACTTTGA 241
DB 185 TGCCAAATTTGTTGAAGCTGATTGATGAATTCACGCCCTTGGAAATPACCGTATCACTTTGA 244
QY 242 ACGGAGATTGATCATGCATTGCAATGATTTATGAACAATATGTTGATGATGATGATGATG 301
DB 245 ACGGAGATTGATCATGCATTGCAATGATTTATGAACAATATGTTGATGATGATGATGATG 304
QY 302 TGACCGCTCTCTTATGTTTCCGTCTTATCGCAAGCAAGGATATTTGTTACATGTGA 361
DB 305 TGACCGCTCTCTTATGTTTCCGTCTTATCGCAAGCAAGGATATTTGTTACATGTGA 364
QY 362 TGTITTCATTAACATATAAGACAAAATGGAGCGTTCAAGCAATCGTTAGCTAATGATGT 421
DB 365 TGTITTCATTAACATATAAGACAAAATGGAGCGTTCAAGCAATCGTTAGCTAATGATGT 424

QY 422 TGAAGCTTTGCTTGTAGTTGTACGAAGCAACTTCTATGAGGCTACCTGGGAGATTATATT 481
DB 425 TGAAGCTTTGCTTGTAGTTGTACGAAGCAACTTCTATGAGGCTACCTGGGAGATTATATT 484
QY 482 AGAAGATGCTTCTTGGTTTATACAGATCTCTGTCTTAGCATTTATGACAAAAGATGCTTTTC 541
DB 485 AGAAGATGCTTCTTGGTTTATACAGATCTCTGTCTTAGCATTTATGACAAAAGATGCTTTTC 544
QY 542 TACAAACCCCGCTCTTTTACCGCAATACAAACGGGCACTAAAGCAACCCCTTTGAAAAG 601
DB 545 TACAAACCCCGCTCTTTTACCGCAATACAAACGGGCACTAAAGCAACCCCTTTGAAAAG 604
QY 602 GTTGCCAAAGATAGAGCGCGCGCAGTACATTCCTTTCTATCAACAAAGATTTCTATAA 661
DB 605 GTTGCCAAAGATAGAGCGCGCGCAGTACATTCCTTTCTATCAACAAAGATTTCTATAA 664
QY 662 CAAGACTTTTACTTAAACTTGTAAAGTTAGAGTTCAATTTGCTTTCAGTCAATTGCAAGA 721
DB 665 CAAGACTTTTACTTAAACTTGTAAAGTTAGAGTTCAATTTGCTTTCAGTCAATTGCAAGA 724
QY 722 AGAGCTCAGCCATGTGTGCAAAATGGTGAAGCTTTTCGATATCAAGAAAGCAACGACCTTG 781
DB 725 AGAGCTCAGCCATGTGTGCAAAATGGTGAAGCTTTTCGATATCAAGAAAGCAACGACCTTG 784
QY 782 TTTAAGAGATAGAAATTTGTAATGTCTTTTGGGACTAGGTTTCAGGCTATGAGCCACA 841
DB 785 TTTAAGAGATAGAAATTTGTAATGTCTTTTGGGACTAGGTTTCAGGCTATGAGCCACA 844
QY 842 GTATTCCCGGGCTAGAGTTTCTTCCAAAGCTGTGCTTTATATACTCTTATAGATGA 901
DB 845 GTATTCCCGGGCTAGAGTTTCTTCCAAAGCTGTGCTTTATATACTCTTATAGATGA 904
QY 902 CACTTATGATGGTATGTGACTTATCAAGAACTTAAGATCTTTACTGAAGCTGTGGAAG 961
DB 905 CACTTATGATGGTATGTGACTTATCAAGAACTTAAGATCTTTACTGAAGCTGTGGAAG 964
QY 962 GTGGTCAATTAATCATGCTTTAGACACATTCGAAATACATGAACCGGATATCAAAATTAT 1021
DB 965 GTGGTCAATTAATCATGCTTTAGACACATTCGAAATACATGAACCGGATATCAAAATTAT 1024
QY 1022 CATGATATACATACAGAAATGGAAAGATTTCTTCAAGAGGAGGAGACAGATCTATT 1081
DB 1025 CATGATATACATACAGAAATGGAAAGATTTCTTCAAGAGGAGGAGACAGATCTATT 1084
QY 1082 TAACTGCGGCAAAAGATTTGTGAAGAGTTTGTAGAAACCTGATGTTGAAAGCAAAATG 1141
DB 1085 TAACTGCGGCAAAAGATTTGTGAAGAGTTTGTGAAGAGCTGATGTTGAAAGCAAAATG 1144
QY 1142 GGCAAATGAGGACACATACCAACCACTGAAGAGCATGATCCAGTTGTAATCATCTATG 1201
DB 1145 GGCAAATGAGGACACATACCAACCACTGAAGAGCATGATCCAGTTGTAATCATCTATG 1204
QY 1202 CGGTGCTAACCTGCTTACAAACCTTTGTTATCTTGGCATGAGTGATATTTCAAAAAGA 1261
DB 1205 CGGTGCTAACCTGCTTACAAACCTTTGTTATCTTGGCATGAGTGATATTTCAAAAAGA 1264
QY 1262 GTCTGTGCAATGGGCTGTCTCTGCACCTCTCTTTTGTAGATCTCAGGTATCTTGTGTCG 1321
DB 1265 GTCTGTGCAATGGGCTGTCTCTGCACCTCTCTCTTTTGTAGTACTCAGGTATCTTGTGTCG 1324
QY 1322 ACGCTTAAATGATCTCATGACCAAGCGCGAGCAAGAAAGAAACATAGTTTCATCGAG 1381
DB 1325 ACGCTTAAATGATCTCATGACCAAGCGCGAGCAAGAAAGAAACATAGTTTCATCGAG 1384
QY 1382 CCTTGAAGTTATATGAAGGAATATATGTCAATGAGGAGTATGCCCAACCTTGTATTGA 1441
DB 1385 CCTTGAAGTTATATGAAGGAATATATGTCAATGAGGAGTATGCCCAACCTTGTATTGA 1444
QY 1442 CAAGAAAGTAGAAGATGTGTGAAAGATATAAACCAGGAGTACCTCAACATAAAACAT 1501
DB 1445 CAAGAAAGTAGAAGATGTGTGAAAGATATAAACCAGGAGTACCTCAACATAAAACAT 1504
QY 1502 TCAAAGCGGCTTATTGATGGCTGTGATCTATTGTTGCCAGCTTTCTTGAAGTTCAATATGC 1561

Db 1505 TCCAGGCCGTTATTGATGCGTGTGATCTATTGTGCCAGTCTCTTTGAAGTTCATATGCG 1564
QY 1562 AGGAAGGATAAATTACACGATATGGGAGAGCAATACAAAATCTCATTAAGTCTCTACT 1621
Db 1565 AGGAAGGATAAATTACACGATATGGGAGAGCAATACAAAATCTCTATAAAGTCTCTACT 1624
QY 1622 CGTTTATCCTATGATATATGAGGATC 1648
Db 1625 CGTTTATCCTATGATATATGACTACC 1651

RESULT 2
BD227440 2076 bp DNA linear PAT 17-JUL-2003
LOCUS Transgenic amorpho-4,11-diene synthesis.
DEFINITION BD227440
ACCESSION BD227440.1 GI:33037210
VERSION JP 2002523101-A/2.
KEYWORDS Artemisia annua (sweet wormwood)
SOURCE Artemisia annua
ORGANISM Artemisia annua

REFERENCE
AUTHORS Wallaart,T.E. and Bouwmeester,H.J.
TITLE Transgenic amorpho-4,11-diene synthesis
JOURNAL Patent: JP 2002523101-A 2 30-JUL-2002;
GENCLIPP BIOTECHNOLOGY BV
COMMENT OS Artemisia annua
PN JP 2002523101-A/2
PD 30-JUL-2002
PF 27-AUG-1999 JP 2000567711
PR 27-AUG-1998 EP 98202854.0
PI THORVALD BELCO WALLAART,HENDRIK JAN BOUWMEESTER PC
C12N15/09,C12N15/09,A01H5/00,C12N1/19,C12N1/21,C12N5/10,C12N9/ PC
88,C12P5/00.
PC C12P17/18//C12N9/04,(C12N1/19,C12R1:84),(C12N1/19,C12R1:645),
PC (C12N9/04,C12R1:91),(C12N9/04,C12R1:865),(C12N9/04,C12R1:84),
PC (C12N9/04,C12R1:19),(C12N9/04,C12R1:645),(C12N9/88,C12R1:91),
PC (C12N9/88,C12R1:19),(C12N9/88,C12R1:84),(C12N9/88,C12R1:865),
PC (C12N9/88,C12R1:645),(C12N15/09,C12R1:91),(C12P5/00,C12R1:19),
PC (C12P5/00,C12R1:84),(C12P5/00,C12R1:91),(C12P5/00,C12R1:865),
PC (C12P5/00,C12R1:645),(C12P17/18,C12R1:645),(C12P17/18,C12R1:19) PC
PC (C12P17/18,C12R1:91),(C12P17/18,C12R1:84),(C12P17/18,C12R1:865) PC
C12N15/00
CC C12N15/00,C12N5/00,(C12N15/00,C12R1:91)
CC Start codon (ARG/Met) starts at nucleic acid position (6) and
CC stop codon
CC (TGA) at position 1644
PH Key Location/Qualifiers
FT CDS Location/Qualifiers
1. 2076
/organism="Artemisia annua"
/mol_type="genomic DNA"
/db_xref="taxon:35608"

FEATURES
source
i. 2076
/organism="Artemisia annua"
/mol_type="genomic DNA"
/db_xref="taxon:35608"

ORIGIN
Query Match 99.3%; Score 1637.4; DB 6; Length 2076;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1641; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 CATGGCCTTACAGAAACCACTATTTCGCCCATTCGCCACTTTCCTCCCAAGCATTTG 61
Db 5 CATGTCTACTACAGAAACCACTATTTCGCCCATTCGCCACTTTCCTCCCAAGCATTTG 64
QY 62 GGGAGATCAGTTTCTCATCTATCAAAACCAAGTAGAGCAAGGGTGGAAACAGATAGTCAA 121

Db 65 GGGAGATCAGTTTCTCATCTATCAAAAGCAAGTAGAGCAAGGGTGGAAACAGATAGTCAA 124
QY 122 TGAATTTAAAAAAGAGTGGGCAACTACTATAAAGAGCTTTTGGATATTTCCATGAACA 181
Db 125 TGAATTTAAAAAAGAGTGGGCAACTACTATAAAGAGCTTTTGGATATTTCCATGAACA 184
QY 182 TGCCTAATTTGTTGAAGCTGATTCGATGAAATTCACAGCGCTTTGGAATACCGTATCACTTTGA 241
Db 185 TGCCTAATTTGTTGAAGCTGATTCGATGAAATTCACAGCGCTTTGGAATACCGTATCACTTTGA 244
QY 242 ACGGAGATTTGATCATGCAATTTGCAATTTATTAAGAAACATATCGTGATATACTGGAATGG 301
Db 245 ACGGAGATTTGATCATGCAATTTGCAATTTATTAAGAAACATATCGTGATATACTGGAATGG 304
QY 302 TGACCGCTCTCTCTTATGTTTCCGCTCTTATGCGAAAGCAAGATATTTATGTTATGATGGA 361
Db 305 TGACCGCTCTCTCTTATGTTTCCGCTCTTATGCGAAAGCAAGATATTTATGTTATGATGGA 364
QY 362 TGTCTTCAATAACTATAAAGACAAATAATGAGCGGTTCAAGCAATCGTTAGCTAATGATGT 421
Db 365 TGTCTTCAATAACTATAAAGACAAATAATGAGCGGTTCAAGCAATCGTTAGCTAATGATGT 424
QY 422 TCAAGGTTTGTGAGTTGTGACGAACAACTTCTATGAGGGTACCTGGGGAGATATATTT 481
Db 425 TGAAGGTTTGTGAGTTGTGACGAACAACTTCTATGAGGGTACCTGGGGAGATATATTT 484
QY 482 AGAAGATGCTCTTGGTTTACACGATCTCGTCTTAGCATTTATGACAAAGATGCTTTTTC 541
Db 485 AGAAGATGCTCTTGGTTTACACGATCTCGTCTTAGCATTTATGACAAAGATGCTTTTTC 544
QY 542 TACAAACCCCGCTCTTTTACCGAAATACAAACGGGCACTAAAGCAACCCCTTTGGAAAG 601
Db 545 TACAAACCCCGCTCTTTTACCGAAATACAAACGGGCACTAAAGCAACCCCTTTGGAAAG 604
QY 602 GTTGCAAGAATAGAGCGCGCGCAGTACATCTCTTCTATCAACAAAGATTCATATAA 661
Db 605 GTTGCAAGAATAGAGCGCGCGCAGTACATCTCTTCTATCAACAAAGATTCATATAA 664
QY 662 CAAGACTTTTACTTAACTTGTCTAAGTTAGAGTTCAATTTGCTTCAGTCATTGCAACGA 721
Db 665 CAAGACTTTTACTTAACTTGTCTAAGTTAGAGTTCAATTTGCTTCAGTCATTGCAACGA 724
QY 722 AGAGCTCAGCCATGTGTGCAAAATGGTGGAAAGCTTTTCGATATCAAGAAAGAACCGACCTTG 781
Db 725 AGAGCTCAGCCATGTGTGCAAAATGGTGGAAAGCTTTTCGATATCAAGAAAGAACCGACCTTG 784
QY 782 TTTAAGAGATAGAAATTTGTAATGTCTATTTTGGGACTAGGTTTCAAGCTATGAGCCACA 841
Db 785 TTTAAGAGATAGGTTTGTGTAATGTCTATTTTGGGACTAGGTTTCAAGCTATGAGCCACA 844
QY 842 GTATTTCCCGGCTAGAGTTTCTTCAAAAGCTGTGCTGTTATATACTCTTTATAGATGA 901
Db 845 GTATTTCCCGGCTAGAGTTTCTTCAAAAGCTGTGCTGTTATATACTCTTTATAGATGA 904
QY 902 CACTTATGATGGTATGGTACTTTATGAAGAACTTAAGATCTTTTACTGAAGCTGTGTAAG 961
Db 905 CACTTATGATGGTATGGTACTTTATGAAGAACTTAAGATCTTTTACTGAAGCTGTGTAAG 964
QY 962 GTGGTCAATTACATGCTTAGACACACTTCAGAAATACATGAACCGATATACAAATTTAT 1021
Db 965 GTGGTCAATTACATGCTTAGACACACTTCAGAAATACATGAACCGATATACAAATTTAT 1024
QY 1022 CATGGATACATACAGAAATGGAAAGTTTGTGAAAGCTGTGTAAGAACTGTGTTGAAGCAAAATG 1081
Db 1025 CATGGATACATACAGAAATGGAAAGTTTGTGAAAGCTGTGTAAGAACTGTGTTGAAGCAAAATG 1084
QY 1082 TAACTGCGGCAAAAGAAATTTGTGAAAGCTGTGTAAGAACTGTGTTGAAGCAAAATG 1141
Db 1085 TAACTGCGGCAAAAGAAATTTGTGAAAGCTGTGTAAGAACTGTGTTGAAGCAAAATG 1144
QY 1142 GCGAAATGAGGACACATACCAACCACTGAAGAGCATGATCCAGTTGTATATCATCTG 1201
Db 1145 GCGAAATGAGGACACATACCAACCACTGAAGAGCATGATCCAGTTGTATATCATCTG 1204

```
QY 1202 CGGTGCTAAACCTGCTTACACCACTTGTATCTTGGCATGAGTATATATTCACAAAGA 1261
Db 1205 CGGTGCTAAACCTGCTTACACCACTTGTATCTTGGCATGAGTATATATTCACAAAGA 1264
QY 1262 GTCTGCGAATGGGCTGCTCTGACCTCTCTTTTAGTACTCAGGTATATCTTGTGTCG 1321
Db 1265 GTCTGCGAATGGGCTGCTCTGACCTCTCTTTTAGTACTCAGGTATATCTTGTGTCG 1324
QY 1322 AGCCCTAAATGATCTCATGACCCACAGGCGGAGCAGAGAAAGAAACATAGTTCATCGAG 1381
Db 1325 AGCCCTAAATGATCTCATGACCCACAGGCGGAGCAGAGAAAGAAACATAGTTCATCGAG 1384
QY 1382 CTTTGAAGCTTATGAAGGATATATGCAATGAGGAGTATGCGCCAAACCTTGATTGA 1441
Db 1385 CTTTGAAGCTTATGAAGGATATATGCAATGAGGAGTATGCGCCAAACCTTGATTGA 1444
QY 1442 CAAGGAAGTAGAAGATGTGTGGAAGATATAAACCGGAGTACTCTCAACATAAAACAT 1501
Db 1445 CAAGGAAGTAGAAGATGTGTGGAAGATATAAACCGGAGTACTCTCAACATAAAACAT 1504
QY 1502 TCCAGGCGGTTATTGATGGCTGTGATCTATTTGTGCCAGTTTCTTGAAGTTCATATGC 1561
Db 1505 TCCAGGCGGTTATTGATGGCTGTGATCTATTTGTGCCAGTTTCTTGAAGTTCATATGC 1564
QY 1562 AGGAAGGATTAACCTTCACAGCTATGGGAGACGAAATACAAACATCTCTATAAAGTCTCTACT 1621
Db 1565 AGGAAGGATTAACCTTCACAGCTATGGGAGACGAAATACAAACATCTCTATAAAGTCTCTACT 1624
QY 1622 CGTTATCTCTAGTATATGAGGATC 1648
Db 1625 CGTTATCTCTAGTATATGACTACC 1651

RESULT 3
AF138959 2080 bp mRNA linear PLN 25-SEP-2000
LOCUS Artemisia annua amorph-4,11-diene synthase mRNA, complete cds.
DEFINITION Artemisia annua
ACCESSION AF138959
VERSION AF138959.1 GI:7381216
KEYWORDS Artemisia annua (sweet wormwood)
ORGANISM Artemisia annua
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Asteroideae;
Anthemideae; Artemisia.
Mercke, P., Bengtsson, M., Bouwmeester, H.J., Posthumus, M.A. and
Brodelius, P.E.
Molecular cloning, expression, and characterization of
amorph-4,11-diene synthase, a key enzyme of artemisinin
biosynthesis in Artemisia annua L
Arch. Biochem. Biophys. 381 (2), 173-180 (2000)
20485153
MEDLINE 11032404
PUBMED 11032404
REFERENCE 2 (bases 1 to 2080)
Mercke, P.E. and Brodelius, P.E.
Direct Submission
AUTHORS Submitted (26-MAR-1999) Plant Biochemistry, Lund University, P.O.
Box 117, Lund S-221 00, Sweden
JOURNAL Box 117, Location/Qualifiers
FEATURES
1..2080
/organism="Artemisia annua"
/mol_type="mRNA"
/db_xref="taxon:35608"
/tissue_type="leaf"
/note="Artemisia annua L."
13..1653
/function="catalyzes the production of amorph-4,11-diene
from farnesyl diphosphate"
/note="sesquiterpene synthase"
/codon_start=1
CDS
```

```
/product="amorph-4,11-diene synthase"
/protein_id="AA061439.1"
/db_xref="GI:7381217"
/translation="MSLITEERPIRIANPPSIWGDQFLIYKQVQEQVEQIVNDLKK
EVRLLEALDIPMKHANLLKIDIEIQRLGIPYHFERIDHALQCIYETGRNWDGIR
SSLMFLMRKQGYVTVCDVFNKYKNGAFQSLANDVEGLLELYEAYGVGEIIL
EDALGFTRSLISIMTKDAPLFTBIQKALQKPLKRLPRIEAAQVIFPFYQDQDS
HNKTLKLALFENLLOSLHKBELSHVCKWKAPDIKKNAPCLDRIVBCYFVGLGSG
YEPY:SRARVPFTKAVAVITLIDDTYDAYGTVEELKIPTAEAVERSITCLDTLPEYMK
PIYKLFMTTYTEBEEFLAKGRTDLFNCGEFVKFVRNLVMEKAWNANGHLPITTEH
DPVLIITGGANLLTTCYLGMSVWASAPPLFRYGLGRRLNDLMTTHA
EQERKSSSLESYKNVNEEYATLIYKEVEDVWKDINREYLTITKNIPLLMAY
IYLQFLEVOVAGKDNFTRMGDEVKHLIKSLIVYPMI"
```

ORIGIN

```
Query Match 99.2%; Score 1635.8; DB 8; Length 2080;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1640; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 CATGCGCACTTACAGAGAAACCTTATTCGCCCATTTGCCAATTTCTCCCAAGCATTTG 61
Db 12 CATGTCACTTACAGAGAAACCTTATTCGCCCATTTGCCAATTTCTCCCAAGCATTTG 71
QY 62 GGGAGATCAGTTTCTCATCTATCAAAAGCAAGTAGAGCAAGGGGTGGAAACAGTAGTGA 121
Db 72 GGGAGATCAGTTTCTCATCTATGAAAGCAAGTAGAGCAAGGGGTGGAAACAGTAGTGA 131
QY 122 TGATTTAAAAAAGAGTGGCGCACTACTAAAGCAAGTCTTGGATATTTCTTATGAACA 181
Db 132 TGATTTAAAAAAGAGTGGCGCACTACTAAAGCAAGTCTTGGATATTTCTTATGAACA 191
QY 182 TGCCAAATTTGTTGAAGCTGATTGATGAATTCAGCGCTTGGAAATACCGTATCACATTGA 241
Db 192 TGCCAAATTTGTTGAAGCTGATTGATGAATTCAGCGCTTGGAAATACCGTATCACATTGA 251
QY 242 ACGGGAGATTCATCATGCAATTCGAATGATTTATGAAACATATGTTGATAAAGTGAATGG 301
Db 252 ACGGGAGATTCATCATGCAATTCGAATGATTTATGAAACATATGTTGATAAAGTGAATGG 311
QY 302 TGACCGCTCTTCCCTTATGTTCCGCTTATGCGAAGCAAGGATATTTGTTACATGTGA 361
Db 312 TGACCGCTCTTCCCTTATGTTCCGCTTATGCGAAGCAAGGATATTTGTTACATGTGA 371
QY 362 TGTTTTCAATAACTATAAAGCAAAATAGGACGCTTCAAGCAATCGTTAGCTAATGATGT 421
Db 372 TGTTTTCAATAACTATAAAGCAAAATAGGACGCTTCAAGCAATCGTTAGCTAATGATGT 431
QY 422 TGAAGTTTGTGTTGATGTGTACGAAGCAACTTCTATGAGGGTACCTGGGGAGATTATTT 481
Db 432 TGAAGTTTGTGTTGATGTGTACGAAGCAACTTCTATGAGGGTACCTGGGGAGATTATTT 491
QY 482 AGRAGATCTCTTGGTTTTACACGATCTCGCTTAGCATTTATGACAAAGATGCTTTTTC 541
Db 492 AGRAGATCTCTTGGTTTTACACGATCTCGCTTAGCATTTATGACAAAGATGCTTTTTC 551
QY 542 TACAAACCCCGCTCTTTTATACGAATACAAAGGCACTTAAAGCAACCCCTTTGGAAAAG 601
Db 552 TACAAACCCCGCTCTTTTATACGAATACAAAGGCACTTAAAGCAACCCCTTTGGAAAAG 611
QY 602 GTTGCAGAGATAGAGGGCGGCGAGTACATTCCTTTCTATCAACAACAAGATTCTCATAA 661
Db 612 GTTGCAGAGATAGAGGGCGGCGAGTACATTCCTTTCTATCAACAACAAGATTCTCATAA 671
QY 662 CAAGACTTACTTAAACTGCTAAGTTAGAGTTCAATTTGCTTCAGTTCATTCGACAGGA 721
Db 672 CAAGACTTACTTAAACTGCTAAGTTAGAGTTCAATTTGCTTCAGTTCATTCGACAGGA 731
QY 722 AGAGCTCAGCCATGTGTCAAAATGTGGAAAGCTTTTCGATATCAAGAAGAACCCACTTG 781
Db 732 AGAGCTCAGCCATGTGTCAAAATGTGGAAAGCTTTTCGATATCAAGAAGAACCCACTTG 791
QY 782 TTATAAGATAGAATTTGTTGAATGCTTTTGGGACTAGGTTGGGCTATGAGGCACA 841
```

Db 792 TTTAAGAGATAGAAATGTTGAATGCTACTTTTGGGAGCTAGGTCAGGCTATGAGCCACA 851
QY 842 GTATTCGGGGCTAGAGTTTCTTCAAAAAGCTGTGCTGCTATTAATACTCTTATAGATGA 901
Db 852 GTATTCGGGGCTAGAGTTTCTTCAAAAAGCTGTGCTGCTATTAATACTCTTATAGATGA 911
QY 902 CACTTATGATCGGTATGCTATGATGAAGAACTTAAAGATCTTTAAGCTGTGGAAG 961
Db 912 CACTTATGATCGGTATGCTATGATGAAGAACTTAAAGATCTTTAAGCTGTGGAAG 971
QY 962 GTGGTCAATTTACATGCTTTAGACACACTTCCAGATACATGAAGAAACCGATATACAAATATT 1021
Db 972 GTGGTCAATTTACATGCTTTAGACACACTTCCAGATACATGAAGAAACCGATATACAAATATT 1031
QY 1022 CATGGATACATACACAAATGGAAGAAATTTCTTCAAAAGGAGGAGAAACAGATCTATT 1081
Db 1032 CATGGATACATACACAAATGGAAGAAATTTCTTCAAAAGGAGGAGAAACAGATCTATT 1091
QY 1082 TAACTCGGCAAGAAATTTGGAAGAGTTTGTAGAAACCTGATGTTGAGCAAAATG 1141
Db 1092 TAACTCGGCAAGAAATTTGGAAGAGTTTGTAGAAACCTGATGTTGAGCAAAATG 1151
QY 1142 GGCAAATGAGGACACATACCAACCACTGGAAGCATGATCCAGTTGTAATCATCTG 1201
Db 1152 GGCAAATGAGGACACATACCAACCACTGGAAGCATGATCCAGTTGTAATCATCTG 1211
QY 1202 CGGTGCTAAGCTGTTACAAACACTTGTATCTTTGGCATGATGATATATTCACAAAGA 1261
Db 1212 CGGTGCTAAGCTGTTACAAACACTTGTATCTTTGGCATGATGATATATTCACAAAGA 1271
QY 1262 GTCTGCGAATGGGCTGCTCTGCAACCTCTCTTTTGTAGATACCTCAGATATCTGTCG 1321
Db 1272 GTCTGCGAATGGGCTGCTCTGCAACCTCTCTTTTGTAGATACCTCAGATATCTGTCG 1331
QY 1322 ACCCTTAATGATCTCATGACCCACAGGCGGCAAGAAAGAAACATATCTTCATCGAG 1381
Db 1332 ACCCTTAATGATCTCATGACCCACAGGCGGCAAGAAAGAAACATATCTTCATCGAG 1391
QY 1382 CTTGAAAGTTATATGAAGAAATATATGTCAATGAGGAGTATGCCCAAACTTGAATTA 1441
Db 1392 CTTGAAAGTTATATGAAGAAATATATGTCAATGAGGAGTATGCCCAAACTTGAATTA 1451
QY 1442 CAAGGAAGTATAGATGTGGAAGATATATAACCGAGAGTACCTCAACATAAAACAT 1501
Db 1452 CAAGGAAGTATAGATGTGGAAGATATATAACCGAGAGTACCTCAACATAAAACAT 1511
QY 1502 TCCAGGCGGTTATGATGCTGTGATCTATTTGTCGACGTTTCTTGAAGTTCAATATGC 1561
Db 1512 TCCAGGCGGTTATGATGCTGTGATCTATTTGTCGACGTTTCTTGAAGTTCAATATGC 1571
QY 1562 AGGAAAGGATAAATTACACGATGGGAGACGAATACAAACATCTCAAAAAGTCTCTACT 1621
Db 1572 AGGAAAGGATAAATTACACGATGGGAGACGAATACAAACATCTCAAAAAGTCTCTACT 1631
QY 1622 CGTTTATCTTATGATATAGAGATC 1648
Db 1632 CGTTTATCTTATGATATAGACTACC 1658

RESULT 4

AA251751 2106 bp mRNA linear PLN 24-NOV-2000
LOCUS Artemisia annua mRNA for amorpha-4,11-diene synthase (kcs12 gene).
DEFINITION AJ251751
ACCESSION Artemisia annua (sweet wormwood)
VERSION AJ251751.1 GI:8546577
KEYWORDS amorpha-4,11-diene synthase; kcs12 gene.
SOURCE Artemisia annua
ORGANISM Artemisia annua
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Asteroidae;
Anthemideae; Artemisia.
1

REFERENCE

AUTHORS Chang, Y.J., Song, S.H., Park, S.H. and Kim, S.U.
TITLE Amorpha-4,11-diene synthase of Artemisia annua: cDNA isolation and bacterial expression of a terpene synthase involved in artemisinin biosynthesis
JOURNAL Arch. Biochem. Biophys. 383 (2), 178-184 (2000)
MEDLINE 21036108
PUBMED 11185551
REFERENCE 2 (bases 1 to 2106)
AUTHORS Chang, Y.J.
TITLE Direct Submision
JOURNAL Submitted (10-DEC-1999) Chang Y.J., Research Center for New Bio-Materials in Agriculture, Seoul National University, Kwonsun-gu, Soedun-dong 103, Suwon 441-744, SOUTH KOREA
FEATURES
source
1. .2106
/organism="Artemisia annua"
/mol_type="mRNA"
/db_xref="taxon:35608"
/country="South Korea:Suwon"
1. .2106
/gene="kcs12"
34. .1675
/gene="kcs12"
/function="sesquiterpene cyclase"
/codon_start=1
/product="amorpha-4,11-diene synthase"
/protein_id="CA894691.1"
/db_xref="GI:8546578"
/db_xref="GOA:O9LW98"
/db_xref="SPTREMBL:Q9LW98"
/translation="MSLTPEEKPIRPINPPSPISWIGDQFLIYKQVEQVQIIVNDLKK
EVRQLKEALDIPMKHANLLKLLDEIRLIGIPYHFEREIDHALQCIYETVYDGNWNGDR
SSLMFLRMKQYVYTCVFNKYKNGAFKQSLANDVEGLLELYEATSRVPOEII
EDALGFTSRSLSMTKDAPSTNPALFTEIQRALKQPLWKLRLPRIEAAQYIPFQQQDS
ENKTLKAKLEFNLLQSLHKEBELSHVCKWKAFLDKKAPCLDRIDRVCEVFWGLGSG
PEPQSRARVFEFTKAVANTLLDDTYDAYGVTEELKFTFAVERWSITCLDITLPEYMK
FYIKGFMDTYTEMEFLKAKEGIDLFNCGKEFKFVRNLMVAKWANEHGIPTTEEH
DPVITGGANLLTTCYLGMSDIPTKSEVNAVASAPPLFRYSGLIGRLRRLLDMTHKA
EQERKXSSNLTYSYKYNVEYATLIYKEVEDYVKDKINREYLTNTKIPRLLMAV
IYLCFLEVQVAGKDNFTRMGDEYKHLIKSLIVYPMISI"
ORIGIN
Query Match 99.1%; Score 1634.2; DB 8; Length 2106;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1639; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 2 CATGCACTTACAGAGAAAACCTATTGCGCCCATTTGCCAATTTCTCTCCACGATTG 61
Db 33 CATGCACTTACAGAGAAAACCTATTGCGCCCATTTGCCAATTTCTCTCCACGATTG 92
QY 62 GGGAGATCAGTTTCTCATCTATCAAAAGCAAGTAGAGCAAGGGGTGGAACAGATAGTAA 121
Db 93 GGGAGATCAGTTTCTCATCTATCAAAAGCAAGTAGAGCAAGGGGTGGAACAGATAGTAA 152
QY 122 TGATTTAAAAAGAGTCGGCACTACTATAAAGAGCTTTTGGATATTCCTATGAACA 181
Db 153 TGATTTAAAAAGAGTCGGCACTACTATAAAGAGCTTTTGGATATTCCTATGAACA 212
QY 182 TGCCCAATTTGTTGAAGCTGATTGATGAATTCACCCCTTTGGAAATCCGCTATCACTTTGA 241
Db 213 TGCCCAATTTGTTGAAGCTGATTGATGAATTCACCCCTTTGGAAATCCGCTATCACTTTGA 272
QY 242 ACGGAGATTTGATCATGCAATTCGAATTTATGAACAATATGGTGATAACTGGAATGG 301
Db 273 ACGGAGATTTGATCATGCAATTCGAATTTATGAACAATATGGTGATAACTGGAATGG 332
QY 302 TGACCGCTCTTCTTATGTTGTTGCTTATGCGAAAGCAAGATATTTATGTTACATGTA 361
Db 333 TGACCGCTCTTCTTATGTTGTTGCTTATGCGAAAGCAAGATATTTATGTTACATGTA 392
QY 362 TCTTTTCAATTAATAAGACAAATATGAGCGGTTCACGAATCGTTAGCTATGATGT 421
Db 393 TCTTTTCAATTAATAAGACAAATATGAGCGGTTCACGAATCGTTAGCTATGATGT 452

QY	422	TGAAGGTTTCGTTGAGTTGTACGAAGCAACTTCTATGAGGGTACCTGGGAGATTATATT	481
DB	453	TGAAGGTTTCGTTGAGTTGTACGAAGCAACTTCTATGAGGGTACCTGGGAGATTATATT	512
QY	482	AGAAGATGCTCTTGGTTTTCACGATCTCGTCTTAGCATATATGACAAAAAGATGCTTTTTC	541
DB	513	AGAAGATGCTCTTGGTTTTCACGATCTCGTCTTAGCATATATGACAAAAAGATGCTTTTTC	572
QY	542	TACAAACCCCGCTCTTTTTCACGAAATACAAACGGGCACTAAGCAACCCCTTTGGAAAAAG	601
DB	573	TACAAACCCCGCTCTTTTTCACGAAATACAAACGGGCACTAAGCAACCCCTTTGGAAAAAG	632
QY	602	GTTGCCAAGAATAGAGCGCGGAGTACATTCCTTTCTATCAACAACAAGATCTCTATAA	661
DB	633	GTTGCCAAGAATAGAGCGCGGAGTACATTCCTTTCTATCAACAACAAGATCTCTATAA	692
QY	662	CAAGACTTTACTTTAAACTTGTCTAAGTTAGAGTTCAATTTGCTTCAGTCAATTGCACAAGGA	721
DB	693	CAAGACTTTACTTTAAACTTGTCTAAGTTAGAGTTCAATTTGCTTCAGTCAATTGCACAAGGA	752
QY	722	AGAGCTCAGCCATGTGTGCAAAATGGTGGAAAGCTTTTCGATATCAAGAAGAACGCACTTG	781
DB	753	AGAGCTCAGCCATGTGTGCAAAATGGTGGAAAGCTTTTCGATATCAAGAAGAACGCACTTG	812
QY	782	TTTAAAGAGATAGAAATTCGTAATGCTACTTTTGGGGACTAGGTTTCAGGCTATGAGCCACA	841
DB	813	TTTAAAGAGATAGAAATTCGTAATGCTACTTTTGGGGACTAGGTTTCAGGCTATGAGCCACA	872
QY	842	GTATTCGGGGCTAGAGTTTCTTCAAAAAGCTGTGCTGTTATATACTCTTATAGATGA	901
DB	873	GTATTCGGGGCTAGAGTTTCTTCAAAAAGCTGTGCTGTTATATACTCTTATAGATGA	932
QY	902	CACTTATGATGCGTATGCTACTTATGAAGAACTTTAAGATCTTTTACTCAAGCTGTGGAAG	961
DB	933	CACTTATGATGCGTATGCTACTTATGAAGAACTTTAAGATCTTTTACTGAACTGTGGAAG	992
QY	962	GTGTCATAATACATGCTTAGACACACTTCCAGAAATACATGAACCCGATATACAAATTTATT	1021
DB	993	GTGTCATAATACATGCTTAGACACACTTCCAGAAATACATGAACCCGATATACAAATTTATT	1052
QY	1022	CATGGATACATACACAGAAATGGNAGAAATTCCTTGCAAAAGGAGGGAAGAACAGATCTATT	1081
DB	1053	CATGGATACATACACAGAAATGGNAGAAATTCCTTGCAAAAGGAGGGAAGAACAGATCTATT	1112
QY	1082	TAACTGCGGCAAAAGAAATTTGTGAAGAAGTTGTTAGAAAACCTGATGTTTGAAGCAAAATG	1141
DB	1113	TAACTGCGGCAAAAGAAATTTGTGAAGAAGTTGTTAGAAAACCTGATGTTTGAAGCAAAATG	1172
QY	1142	GGCAATGAGGACACATACAAACCATGAAGAGCATGATCCAGTTGTGTAATCTACTGCG	1201
DB	1173	GGCAATGAGGACACATACAAACCATGAAGAGCATGATCCAGTTGTGTAATCTACTGCG	1232
QY	1202	CGGTGCTAACTCGTTTCAACAACCTTGTTATCTTGGCATGAGTGATATATTCACAAAAGA	1261
DB	1233	CGGTGCTAACTCGTTTCAACAACCTTGTTATCTTGGCATGAGTGATATATTCACAAAAGA	1292
QY	1262	GTCTGTGCAATGGGCTGCTCTGCACTCCCTCTTTTGTAGTACTCAGGTATACTTGGTCG	1321
DB	1293	GTCTGTGCAATGGGCTGCTCTGCACTCCCTCTTTTGTAGTACTCAGGTATACTTGGTCG	1352
QY	1322	ACGCCTAAATGATCTCATGACCCCAAGCGCGGAGCAAGAAAGAAAACATAGTTCATCCAG	1381
DB	1353	ACGCCTAAATGATCTCATGACCCCAAGCGCGGAGCAAGAAAGAAAACATAGTTCATCCAG	1412
QY	1382	CCTTGAAAGTTATATGAAGGAATATAATGTCAATGAGGAGTATGCCCAAAACCTTGATTTA	1441
DB	1413	CCTTGAAAGTTATATGAAGGAATATAATGTCAATGAGGAGTATGCCCAAAACCTTGATTTA	1472
QY	1442	CAAGGAAGTGAAGATGTGTGGAAGATATAAACCAGAGATGACCTCAACTCAAAACAT	1501
DB	1473	CAAGGAAGTGAAGATGTGTGGAAGATATAAACCAGAGATGACCTCAACTCAAAACAT	1532

Qy	1502	TCCAAGCGCGTTATTCATGGCTCGTGATCTATTTGTCGAGTTTCTTGAAGTTCAAATATGC	1561
Db	1533	TCCAAGCGCGTTATTCATGGCTCGTGATCTATTTGTCGAGTTTCTTGAAGTTCAAATATGC	1592
Qy	1562	AGGAAGGATAAATTCCACAGTATGGGAGACGAATCAAAACATCTCATAAAGTCTCTTACT	1621
Db	1593	AGGAAGGATAAATTCCACAGTATGGGAGACGAATCAAAACATCTCATAAAGTCTCTTACT	1652
Qy	1622	CGTTTATCCTTATGAGTATATGAGGATC	1648
Db	1653	CGTTTATCCTTATGAGTATATGAGTATGATAC	1679

[illegible]

REFERENCE 1 (bases 1 to 2067)
AUTHORS Liu, Y., Ye, H.C. and Li, G.F.
TITLE Cloning of sesquiterpene cyclase cDNA from Artemisia annua
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2067)
AUTHORS Liu, Y., Ye, H.C. and Li, G.F.
TITLE Direct Submission
JOURNAL Submitted (08-DEC-2000) Department of Cell and Gene Engineering,

FEATURES

source

gene

CDS

1. .2067

/organism="Artemisia annua"

/mol_type="mRNA"

/strain="001"

/db_xref="taxon.35608"

/tissue_type="leaf"

/country="China: Sichuan"

1. .2067

/gene="AMS1"

13. 1653

/gene="AMS1"

/note="sesquiterpene cyclase"

/codon_start=1

/product="amorpha-4,11-diene synthase"

/protein_id="AAK15696.1"

/db_xref="GI:13195432"

/translations="MLTSEKIRPIANTPPPSIWGQDFLIYKQVEGGVEQIVNDLKK
EVLQLEALDIPMKHANLLKIDETQRLGIPYHFEQIDHALQCIIYETGDNWGDGR
SYLWFLRMKRGQYVTFNNYKQDKGAFKOSLANDVEGLLETPKSNRVRGEIENR
EDALGTGRSLSTMTKQZSTNPALSTEQRALKQPLKRLPRIEAQIIPFYQOQDS
HNKTLIAKLLENLQSLKHELSHVCKWKAEDIKKNAPCLDRIVCYFYWGLSGG
FEPQYSRARVFTKAVITLDDTDYAGTYEELKIFTEAVERNSITGLDITLPEYMK
PIYKLFMDYITEMEELKGRDRLDFNCGEFVKFEFRLMVEAKNWAHGHIFTEKH
DPVITITGAGNLLTTCYLGMDITPKESVEWAVSAPPLFRYSQGLGRINDLMTTSH
EQBRKSSLSLISYMEKYNVNSYEAQTLIKVEVEDWKRQDINREYLTATKNIPRLLMA
IYVLCOFLEVOYACQNDFTNMGDEYKHLIKSLIYVPMVSI"

Query Match	98.5%;	Score 1626.2;	DB 8;	Length 2067;
Best Local Similarity	99.3%;	Pred. NO. 0;		
Matches 1634;	Conservative 0;	Mismatches 13;	Indels 0;	Gaps 0;

Qy	2	CAATGGCACTTACAGAGAAAAA	CCTATTGCGCCCAATTGCCA	CTTTCTCTCCACGCAATTG	61
Db	12	CAATGGCACTTACAGAGAAAAA <th>CCTATTGCGCCCAATTGCCA</th> <th>CTTTCTCTCCACGCAATTG</th> <th>71</th>	CCTATTGCGCCCAATTGCCA	CTTTCTCTCCACGCAATTG	71

QY 62 GGAGATCAGTTTCTCATCTATCAAAAGCAAGTAGAGCAAGGGGTGGAAACAGATAGTAA 121
Db 72 GGAGATCAGTTTCTCATCTATGAAGCAAGTAGAGCAAGGGGTGGAAACAGATAGTAA 131
QY 122 TGATTTAAAAAGAGAGTGGCAACTACTAAAGAAAGCTTTGGATATTCCTATGAACA 181
Db 132 TGATTTAAAAAGAGAGTGGCAACTACTAAAGAAAGCTTTGGATATTCCTATGAACA 191
QY 182 TGCAATTTGTTGAAGCTGATGATGAATCAACGCTTGGATACCGTATCACTTTGA 241
Db 192 TGCAATTTGTTGAAGCTGATGATGAATCAACGCTTGGATACCGTATCACTTTGA 251
QY 242 ACCGGAGATTGATCATGCATTGCAATGATTTATGAACAATATGATGATGATGATG 301
Db 252 ACAGGAGATTGATCATGCATTGCAATGATTTATGAACAATATGATGATGATGATG 311
QY 302 TGACCGCTCTTCCCTTATGTTTCCGCTTATGCGAAAGCAAGATATTTATGATGATG 361
Db 312 TGACCGCTCTTCCCTTATGTTTCCGCTTATGCGAAAGCAAGATATTTATGATGATG 371
QY 362 TGTGTTCAATAATCAATAAGCAAAATGGAGCGTTCAAGCAATCGTTAGCTAATGATG 421
Db 372 TGTGTTCAATAATCAATAAGCAAAATGGAGCGTTCAAGCAATCGTTAGCTAATGATG 431
QY 422 TGAAGGTTTGTGAGTTGACGAAGCAACTTCTATGAGGTTACCTGGGAGATTAAT 481
Db 432 TGAAGGTTTGTGAGTTGACGAAGCAACTTCTATGAGGTTACCTGGGAGATTAAT 491
QY 482 AGAAGATGCTCTTGGTTTACAGATCTGCTTTAGCAATTTATGACAAAGATGCTTTTC 541
Db 492 AGAAGATGCTCTTGGTTTACAGATCTGCTTTAGCAATTTATGACAAAGATGCTTTTC 551
QY 542 TACAAACCCCGCTCTTTTACCGAAATCAACGGGCACTTAAGCAACCCCTTTGAAAG 601
Db 552 TACAAACCCCGCTCTTTTACCGAAATCAACGGGCACTTAAGCAACCCCTTTGAAAG 611
QY 602 GTTGCCAAAGATAGAGGGGGGAGTCAATCTCTTTCTATCAACCAAGATTTCTATAA 661
Db 612 GTTGCCAAAGATAGAGGGGGGAGTCAATCTCTTTCTATCAACCAAGATTTCTATAA 671
QY 662 CAAGACTTTACTTAACTTGTAACTTGAAGTCAATTTGCTTCACTTCACTTCACTCA 721
Db 672 CAAGACTTTACTTAACTTGTAACTTGAAGTCAATTTGCTTCACTTCACTTCACTCA 731
QY 722 AGAGTCAAGATGCTGCAATGCTGGAAGCTTTGATATCAAGAAAGCAAGCTTTG 781
Db 732 AGAGTCAAGATGCTGCAATGCTGGAAGCTTTGATATCAAGAAAGCAAGCTTTG 791
QY 782 TTTAAGATAGAAATGTTGAACTGCTATTTTGGGAGCTAGGTTCAAGGCTATGAGCCACA 841
Db 792 TTTAAGATAGAAATGTTGAACTGCTATTTTGGGAGCTAGGTTCAAGGCTATGAGCCACA 851
QY 842 GTATTTCCCGGCTAGAGTTTCTTCAAAAGCTGTTGCTGTTTAACTTTATAGATGA 901
Db 852 GTATTTCCCGGCTAGAGTTTCTTCAAAAGCTGTTGCTGTTTAACTTTATAGATGA 911
QY 902 CACTTATGATGCTGATGCTATGATGAACTTAAAGCTTTTACTGAAAGCTTTGAAAG 961
Db 912 CACTTATGATGCTGATGCTATGATGAACTTAAAGCTTTTACTGAAAGCTTTGAAAG 971
QY 962 GTGGTCAATTACATGCTTACACACTTCCAGATATCAAGAAACCGATATACAAATTA 1021
Db 972 GTGGTCAATTACATGCTTACACACTTCCAGATATCAAGAAACCGATATACAAATTA 1031
QY 1022 CATGATATACATACAGAAATGGAAGATTTCTTCAAAAGGAGGGAAGACAGATCTATT 1081
Db 1032 CATGATATACATACAGAAATGGAAGATTTCTTCAAAAGGAGGGAAGACAGATCTATT 1091
QY 1082 TAACTGCGCAAGAAATTTGTGAAGAGTTTGTAGAAACCTGATGTTGAACCAAAATG 1141
Db 1092 TAACTGCGCAAGAAATTTGTGAAGAGTTTGTAGAAACCTGATGTTGAACCAAAATG 1151

QY 1142 GGCAATGAGGGACACATACCAACCACTGAAGAGCATGATCCAGTTGTATCTATTCTG 1201
Db 1152 GGCAATGAGGGACACATACCAACCACTGAAGAGCATGATCCAGTTGTATCTATTCTG 1211
QY 1202 CGGTCTTAACCTGCTTACAAACAACTTGTATCTTGGCATGAGTGAATATTCACAAAGA 1261
Db 1212 CGGTCTTAACCTGCTTACAAACAACTTGTATCTTGGCATGAGTGAATATTCACAAAGA 1271
QY 1262 GTCTGTGGAATCGGCTGCTCTGCACTCTCTTTTGTAGATCTCAGGTATCTTGGTCTG 1321
Db 1272 GTCTGTGGAATCGGCTGCTCTGCACTCTCTTTTGTAGATCTCAGGTATCTTGGTCTG 1331
QY 1322 AGCCCTTAATGATCTCATGACCAACCAAGCCGAGCAAGAAAGAAACATAGTTCTATCGAG 1381
Db 1332 AGCCCTTAATGATCTCATGACCAACCAAGCCGAGCAAGAAAGAAACATAGTTCTATCGAG 1391
QY 1382 CTTTCAAAAGCTTATGAAGGAATATATGTCATGAGGAGTATGCCCAAACTTTGATT 1441
Db 1392 CTTTCAAAAGCTTATGAAGGAATATATGTCATGAGGAGTATGCCCAAACTTTGATT 1451
QY 1442 CAAGGAAGTAGAAGTGTGTGGAAGATATAAACCAGAGTACCTCACTCACTAAACAT 1501
Db 1452 CAAGGAAGTAGAAGTGTGTGGAAGATATAAACCAGAGTACCTCACTCACTAAACAT 1511
QY 1502 TCCAGGCCGCTTATGATGCTGCTGATCTATTTGTGCGAGTTCTTGAAGTTCAATATGC 1561
Db 1512 TCCAGGCCGCTTATGATGCTGCTGATCTATTTGTGCGAGTTCTTGAAGTTCAATATGC 1571
QY 1562 AGGAAAGGATACTTCCACAGTATGGAGACGAATACAAACATCTCTATAAGTCTCTACT 1621
Db 1572 AGGAAAGGATACTTCCACAGTATGGAGACGAATACAAACATCTCTATAAGTCTCTACT 1631
QY 1622 CGTTTATCTCTATGATATATGAGGATC 1648
Db 1632 CGTTTATCTCTATGATATATGACTACC 1658

RESULT 6
AF472361 1902 bp mRNA linear PLN 03-JAN-2003
LOCUS Artemisia annua beta-caryophyllene synthase QHS1 (QHS1) mRNA,
DEFINITION complete cds.
ACCESSION AF472361
VERSION AF472361.1 GI:18766933
KEYWORDS
SOURCE Artemisia annua (sweet wormwood)
ORGANISM Artemisia annua
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Asteroideae;
Anthemideae; Artemisia.
REFERENCE
1 (bases 1 to 1902)
AUTHORS Cai, Y., Jia, J.-W., Crock, J., Lin, Z.-X., Chen, X.-Y. and Croteau, R.
TITLE A cDNA clone for beta-caryophyllene synthase from Artemisia annua
JOURNAL Phytochemistry 61 (5), 523-529 (2002)
MEDLINE 22297925
PUBMED 12409018
REFERENCE
2 (bases 1 to 1902)
AUTHORS Cai, Y., Crock, J., Jia, J.-W., Lin, Z.-X., Croteau, R. and Chen, X.-Y.
TITLE Direct Submission
JOURNAL Submitted (20-JAN-2002) National Laboratory of Plant Molecular
Genetics, Institute of Plant Physiology and Ecology, Shanghai
Institutes for Biological Sciences, Chinese Academy of Sciences,
300 Fenglin Road, Shanghai 200032, P. R. China
FEATURES
source Location/Qualifiers
1..1902
/organism="Artemisia annua"
/mol_type="mRNA"
/db_xref="taxon:35608"
gene 1..1902
/gene="QHS1"
CDS 24..1670
/gene="QHS1"

/note="sesquiterpene cyclase"
/codon_start=1
/product="beta-caryophyllene synthase QHS1"
/protein_id="AAL79181.1"
/db_xref="gi:18766934"
/translation="MSVKEEKVIRPIVHPFPPSWADQFLIFDDKQAEQANQVQVNEEL
REVDKLVSLDVOTETNLLKLDIAQLGIAVHPPEEIEQALQHIYDVTGDDWKG
RPSLWPLILQQQFVYVSDIFKNYKXSDGSKESLNDVDELLELYEATYLRVQEG
VLDALVPTRCLEKIALKDLVHTNLTSTYIOERALKDPUHKLRLTLEAURYLPNTEQQ
ASHNESLKLAKLFLNLSQSLKESLSEVSRWMLDVPNNLPYARDRVECYFWALG
YVFPKYQARLFLAKVLSLATVLDYDAYGTBELKIFTEAIQRWSITCDMLPEY
LKLIVQGVLDVYIEMEEIMWGEKKAHLISYAKESKPEFIRSYMEAKWANEGYVPTAQ
EHMSVAFVSSYGLMATTFCVGMGDIIVTDEAFKWLTKPPIIKASCATARLMDDDTHSQ
KEEKERLHVASVSVMKYOVTEEHVLKVENKJEDAKDITRESLVRKIDPMPILMM
RVINLAQWMDVLYKHKGDTNWGSELKDHKISLLVHPPIPI"

ORIGIN

Query Match 52.0%; Score 857; DB 8; Length 1902;
Best Local Similarity 70.8%; Pred. No. 1.4e-186;
Matches 1168; Conservative 0; Mismatches 475; Indels 6; Gaps 2;
2 CATGGCACTTACAGAAAGAAACCTATTTCGCCCCATTTGCGCCATCTTCTCCCAAGCAATTG 61
23 CATGCTCTTAAAGAGAGAAAGTAATTCGGGCCCATTTGCCACATTTCTCTCCCTAGCGTTG 82
62 GGGAGATCAGTCTCTCATCT---ATCAAAGCAAGTAGAGCAAG---GGGTGGAACAGAT 115
83 GCGACACAGTCTCTCATCTTCGATGACAAAGCAAGCAAGCAAGCAAGTGTAGAACAAAGT 142
116 AGTGAATGATTTAAAAAAGAGTGGCGCAACTACTTAAAGAGGCTTTGGATATTCCTAT 175
143 AGTCACAGTTAAGAGAAGATGTGAGAAAGATTTAGTGTCTATCTTTAGATGTTCAAC 202
176 GAAACATGCCAAATTTGTTGAAGCTGATGATGAATGAATCAAGCCCTTGGAAATCCGGTATCA 235
203 GGAGCACACAAATTTGCTAAACCTTATTGATGCTATCCAAAGCTTTGGTATAGCGGTATCA 262
236 CTTTGAACGGGAGATGTGATGATGCAATGCAATGTTATGAAACATATGTTGATGATG 295
263 TTTTGAAGAGAGATGTAGAGCAAGCTTACAAATATTTATGATACATATGTTGATGATG 322
296 GAATGGTGACCGCTCTCCCTTATGTTGCTGCTTATGCGAAAGCAAGGATATTTATGTTAC 355
323 GAAAGCGAGAGCCCTTCCCTTGGTTTTCGAATCTCTCGGCAACAGGTTTTTATGTTTC 382
356 ATGTGATGTTTCAATAACTATAAGACAAATAATGGAGGCTTCAAGCAATCGTTAGCTAA 415
383 TTGTGATATTTTAAAAAATAAAGAGAGATGGTTCAATTAAGGAATCCCTCACCAA 442
416 TGATGTTGAAGGTTGCTTGTAGTCTGACGAACACTTCTATGAGGTACCTGGGAGAT 475
443 CGATGTAAGGCTTGTCTGTAGCTGTATGAGCGACATATTTGAGAGTCAAGGGGAAG 502
476 TATATTAGAAGATGCTCTTGGTTTATACAGATCTCGTCTTAGCATTTATGACAAAAGATGC 535
503 GGTCTAGATGATGCTCTTGTCTTTCACAGGACTTGTCTTGAGAAAATAGCAAGATCT 562
536 TTTTCTTCAAAACCCCGCTCTTTTACCGAAATACAAAGCGGCACTAAAGCAACCCCTTTG 595
563 TGTTCAACCAACCAACTATCTACCTACATACAAAGAGCACTAAAGACAGCCGTTACA 622
596 GAAAGGTTGCCAAGATAGAGGGCGGCGAGTACATTCCTTCTATCAACACAGATTC 655
623 TAAAGGTTGACAGACTAGAGGCAATGCGTTACATTCCTATGTAGCAACACAGATC 682
656 TCATAACAAGACTTTTACTTAACTTGTAAAGTTAGAGTTCAATTTGCTTTCAGTCAATTGCA 715
683 TCATAATGAGTCTTTACTTCAACTTGGAGAGTAGGTTTCACTTACTTCAATCATTTGCA 742
716 CAAGGAAGAGCTCAGCCATGTGTGCAATGTGTGGAAGCTTTCCGATATCAAGAAAGCC 775
743 TAGAAAGAGGCTTAGCGAAGTTTCCAGGTTGTGGAAAGGCTTTGATGTGCCCAACAACTCT 802

QY 776 ACCTTCTTTAAGAGATAGAAATTTGTAATGCTACTCTTTGGGAGCTAGGTTCCAGGCTATGA 835
DB 803 ACCTTATGCAAGAGATAGAAATGTTGTAATGTTATTTTGGCAGCTAGGTGCTATTTTGA 862
QY 836 GCCACAGTATTCOCGGGCTAGAGTTTCTTCCACAAAAGCTGTTGCTGTATATAACTCTTAT 895
DB 863 GCCAAAATATTCCTCAAGCTAGGATCTTTTATAGCAAAAAGTAATTTCCGCTAGCAACTGTTCT 922
QY 896 AGATGACACTTATGATGCGTATGCTATGATGAAGAACTTTAAGATCTTTTACTGAAAGCTGT 955
DB 923 TGACGACACTTATGATGCTTATGAAACCTATGAAAGAACTTTAAGATCTTTTACTGAAAGCAAT 982
QY 956 TGAAGAGTGTCAATTTACATGCTTTAGACACACTTTCCAGAAATACATGAACCGATATACAA 1015
DB 983 TCAAGAGTGTGCAATTTACATGCTATGATATGCTTCCAGAAATACCTGAAACTATTAATCA 1042
QY 1016 ATTATTCATGGATACATACACAGAAATGGAAGAAATTTCTTCAAAAGGAGGAAAGACAGA 1075
DB 1043 AGGAGTCTTGGATATATACATAGAAATGGAAGAAATTAATGGAAAAGAGGAAAGACACA 1102
QY 1076 TCTATTTAACTGCGGCAAGAAATTTGTGAAGAGTTTGTGAAGAACTTTGTAGAAACCTGATGTTGAAGC 1135
DB 1103 TCATCTTAGCTATGCTTAAAGAAATCTATGAAGAGTTTATTAGAAGCTATATGATGGAAGC 1162
QY 1136 AAAATGGGCAATGAGGGACACATACCAACCACTGAAGAGCATGATCCAGTTGTATCAT 1195
DB 1163 AAAATGGGCAATGAGGGGATGTACCAACAGCAGAGGCAATGATGTCAGTTGCTGTTTCT 1222
QY 1196 TACTGCGGTGCTAACTGCTTACCAACACTTGTATCTTGGCATGAGTGTATATTCAC 1255
DB 1223 AAGCAGCGGATATAGCATGCTTGCACCAACATGTTTGTGGCATGGTGATATCGTTAC 1282
QY 1256 AAAAGAGTCTGTCGAATGGGCTGCTCTGACACCTCTCTTTTGTAGATCTCAGGTATACT 1315
DB 1283 AGATGAAGCAATTCAAAATGGGCTCTGACAAAACCTCTCATCAAGAGCTTCATGTGCAAT 1342
QY 1316 TGTGTCAGCGCTAAATGATCTCATGACCCACAGCGCGAGCAAGAAAGAAACATAGTTC 1375
DB 1343 TGCTAGACTCATGATGATATCTACTCTCAAGAGGAGAGAAAGAAATACATGTTGC 1402
QY 1376 ATCGAGCCTTTGAAGTTTATGAAAGGAATATAATGTCAATGAGAGTATGCCAAACCTT 1435
DB 1403 ATCTAGTGTTTGAAGTTTACATGAAGCAATATGATGTGACAGAGGAGCATGTCTTAAAGT 1462
QY 1436 GATTTACAGGAAGTAGAAGATGTGTGAAAGATATATAACCGAGAGTACCTCACAACTAA 1495
DB 1463 ATTAAACAGAAATCGAGGATCGTGGAAAGATATAACCCGAGATCCCTTGTGCGTAA 1522
QY 1496 AAACATTTCCAAAGCCGTTATTTGATGGCTGTGATCTATTTGGCCAGTTCCTTGAAGTTCA 1555
DB 1523 AGATTTCCGATCGCTCTGATGATGCGAGTGTAACTTTGGCACAGGTGATGATGTTT 1582
QY 1556 ATATGCAAGGAAGGATAAATTCACGATGCGGAGACGAATATACAAACATCTCATAAAGTC 1615
DB 1583 ATATAACATAAAGACGGTTTTCATATGTTGGGAGAAAGAACTCAAGGATCATATTTAAATC 1642
QY 1616 TCTACTGTTTATCTTATGATATATGAG 1644
DB 1643 TTTGCTCGTTTCATCTTATACCTATATTAAG 1671

RESULT 7
AAN271793
LOCUS
DEFINITION
Artemisia annua mRNA for putative sesquiterpene cyclase (cASC34 gene).
ACCESSION
AJ271793
VERSION
1
KEYWORDS
cASC34 gene; sesquiterpene cyclase.
SOURCE
Artemisia annua
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

AAN271793 2026 bp mRNA linear PLN 04-OCT-2000
Artemisia annua mRNA for putative sesquiterpene cyclase (cASC34 gene).

398	TTGTGATATATTTTAAAAACTATAAAAAAGAGATGGTTCATTTTAAAGAAATCCCTCAC	457
413	TAATGATGTTGAAAGTTTGCCTTGAGTTGTAAGAAGCAACTTCTATGAGGGTACCTGGGGA	472
458	CAACGATGTAAGAGCTTGTCTGAGCTGTATGAGCGCAGATATTTGAGATGTCAGAGGGA	517
473	GATTATATTGAAGATGCTCTTGGTTTATACAGATCTCGTCTTAGCATTTATGACAAAAGA	532
518	AGGGGTTCTGATGATGCTCTGTGTTTTACAAAGGACTTGTCTTGAGAAAATAGCAAGA	577
533	TGCTTTTTTCTACAAACCCCGCTCTTTTTTACCGAAATACAAAGGGCACTAAAGCAACCCCT	592
578	TCTTGTTCACACCAACCCCAACACTATCTACCTACATACAAGAAGCACTAAACACGCGTT	637
593	TTGAAAAGTTGCCAAGAAATAGAGCGCGCGAGTACATTCCTTCTATCAACACAGA	652
638	ACAAAAAGTTGGCAGACTAGAGGCATTCGATTTACATTCCTATGTAGCAACCAAGC	697
653	TTCTCATACAAAGACTTTTACTTTAACTTCTTAAGTTAGAGTTTCAATTTGCTTCAGTCATT	712
698	ATCTCACAAATGAGTCTTTTACTCAAACTTGGGAAGCTAGGTTTCAACTTACTTCAATCAT	757
713	GCACAAGGAAGAGCTCAGCCATGTGTGCAAAATGGTGGAAAGCTTTTCATATCAAGAAGAA	772
758	GCATAGAAGAGGCTTAACGAAGTTTCAGGTGGTGGAAAGGTCTTGATGCCAACAA	817
773	CGCACTTGTTTAAGAGATAGAAATTTGTTGAATGCTTACTTTTGGGAGCTAGGTTACGGCTA	832
818	TCTACTCTATGCAAGAGATAGAAATGGTTGAATGTTTATTTTGGGCACCTAGGTGTCTATT	877
833	TTGAGCAGAGTATCCCGGGCTAGAGTTTCTTCCAAAAGCTGTGCTGTTTATAACTCT	892
878	TGAGCCAAATATCTCAAGCTAGGATCTTTTTAGCAAAAGTAATTTCCCTGACCAACTGT	937
893	TATGATGACACTTATGATGCGTATGGTACTTTATGAAGAACTTAAGATCTTTTACTGAAGC	952
938	TCCTGACGACACTTATGATGCTTATGGAACCTATGAGAACTTAAGATCTTTTACTGAAGC	997
953	TGTTGAAAGGTGGTCAAATTAATGCTTTAGACACATCTCCAGAAATACATGAAGACCGATATA	1012
998	AAATCAAAGGTGGTCCGTTACATGCATAGATATGCTTCCAGAATACCTGAAAATATTATA	1057
1013	CAAAATATTCATGATACATACACAGAATGGAAGAAATTTCTGCAAAAGGAGGGAAGAC	1072
1058	CCAGGAGTCTTTGGATATATACATAGAATGGAAGAAATATGGGAANAAGAGGHAAGC	1117
1073	AGATCTTATTAACTCGGCGAAAGAAATTTGTGAAGAGTTTGTAGAAAACCTGATGGTTGA	1132
1118	ACATCATCTTACTATGCTTAAGAACTCTATGAAGAGTTTATAGAACTATATGATGA	1177
1133	AGCAAAATGGGCAATGAGGGAACATACCAACCACTGAAGAGCATGATCCAGTTGCTAAT	1192
1178	AGCAAAATGGGCAATGAGGCGTATGTTACCAACAGCAGAGGAGCATATCTCAGTTCCGTT	1237
1193	CATTACTGGCGGTGCTAACCTGCTTACAACTTGTATCTTGGCATGAGTGATATATT	1252
1238	TGTAAGACGCGGTATAGCATGCTTTGCCAACATGTTTTTGGCATGGGTGATATCGT	1297
1253	CACAAAAGATGTGTCGAATGGCTGTCTCTCCACTCTCTTTTATAGTACTTCAGGTAT	1312
1298	TACAGATGAAGCATTCAAATGGGCTCTGCAAAAACCTCTATCATCAAAGCTTCATGTC	1357
1313	ACTTGGTCCAGCCCTAAATGATCTCATGACCCCAAGCGCGAGCAAGAAAGAAAACATAG	1372
1358	AAATGCTAGACTCATGGATGATTTTCACTCTCAAAAGGAGGAGAAAGAAATACATGT	1417
1373	TTTCATCGAGCCTTGAAAGTTTATAGGAAGATATATGTCTCAATGAGAGGTATGCCCAAC	1432
1418	TGCATCTAGTGTGTAAGTTTACATGAAGCAATATGATGTGACAGAGGAGCATGTCCTTAA	1477
1433	CTTGATTTTCAAGGAGTAGAAGATGTGTGGAAAGATATAAAACCGAGAGTACTCTACAAC	1492

Db 1478 AGTATTAAACAGAAATCAGGATCGGTGGAAAGATATACCCGAGAAATCCCTTGTGCG 1537

QY 1493 TAAACACATCCAGCGCGTATTGATGGCTGTGATCTATTGTCAGATTTCTTGAAGT 1552

Db 1538 TAAAGATATCCGATGCTCTGATGATGCGAGTGATTAACCTGGCAGAGTGATGATGT 1597

QY 1553 TCAATATGACGAGAAAGATAACTTCAACATGATGGAGAGAGAAATACAAACATCTCATAAA 1612

Db 1598 TTTATATAACATAAGACCGGTTTCACTAATGTTGGAGAGAACTCAAGATCATATATA 1657

QY 1613 GTCTCTACTGTTTATCTCTAATGATATATAGAG 1644

Db 1658 ATCTTTGCTCGTTTCATCCTATACCTATATAAG 1689

RESULT 8

SCA304452

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

1 gene; germacrene A synthase.

Solidago canadensis

Solidago canadensis

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

asterids; campanulids; Asterales; Asteraceae; Asteroideae;

Asteraceae; Solidago.

1

Prosser, I., Phillips, A.L., Gittings, S., Lewis, M.J., Hooper, A.M.,

Pickett, J.A. and Beale, M.H.

(+)-(10R)-Germacrene A synthase from goldenrod, Solidago

canadensis; cDNA isolation, bacterial expression and functional

analysis

Phytochemistry 60 (7), 691-702 (2002)

22126967

12127586

2 (bases 1 to 1976)

Prosser, I.M.

Direct Submission

Submitted (11-DEC-2000) Prosser I.M., Plant Sciences, IACR-Long

Ashton Research Station, Long Ashton, BS41 9AF, UNITED KINGDOM

Location/Qualifiers

1. .1976

/organism="Solidago canadensis"

/mol_type="rRNA"

/cultivar="Gold Baby"

/db_xref="taxon:59297"

/clone="SC1"

/tissue_type="leaf"

1. .1976

/gene="1"

/gene="1"

41. .1690

/function="conversion of farnesyl diphosphate (FPP) to

(+)-germacrene A"

/codon_start=1

/evidence=experimental

/product="germacrene A synthase"

/protein_id="CAC36896.1"

/db_xref="GI:13624778"

/db_xref="GOA:Q9AR67"

/db_xref="SPTREMBL:Q9AR67"

/translation="MAAKQVEIRPVANVHPSLWQDFLHYDQDEHVEVDQIELL

KEETREILASDDPTKHNLLKLDIVIRGLGIAYFEBHEITQALDHIYSYVDGWNQ

GRTSLWFLRLQCGFYVSDIFNIYKLDNGSPKSLTKDIECMLEIYEAAYMRVQGEI

ILDEALEFTKLEHIAKPLRCNNTLSRHIHEALRPVKRLPRLDALRYIPFYEQQ

DSHNKSLRLAKLGNRLQSLKSKLQSLKSKLQSLKSKLQSLKSKLQSLKSKLQSL

VYFEQYSRSLFLTKTIKMAALLDDYDIYGYVEELFTKAVQKSLTICMTLDY

MKVIYKSLDDVYEMEEIEIKQKQAYQVHYAKESMDLVTSTYMEAKWLHEGHVPD

EHNSVNTIYGGVYKTFASFGVGHGDIYVQESPKWNNPPLIKASD.SRINWDIVG

HKEQQRKHIAKSVENMYKSNLAEDYDFLKERVEDAKWIDNRBTLCIKDILHMK

MPPIINLARVMDMLYKNGDNLKXVQGEIQDYMKSCFINPMSV"

ORIGIN

Query Match

Best Local Similarity

Matches 1119; Conservative

47.11%; Score 776.2; DB 8; Length 1976;

67.8%; Pred. No. 5.5e-168;

Mismatches 523; Indels 9; Gaps 2;

2 CATGCACTTACAGAGAAACCTATTTCGCCCACTTTCCTCCTCAAGCATTTG 61

40 CATGCTCTTAAACAGATGAGGTTATTTCGCCCACTTTCCTCCTCAAGCATTTG 99

62 GGGAGATCAGTTTCTCATCTATCAAAAGCAAGTAGAGCA-----AGGGTGGAAACAGAT 115

100 GGGAGATCAGTTTCTCATCTATGAGCAAGGAGGATGAGCAGGTTGAGGTAGATCAACA 159

116 AGTGAATGATTTTAAAAAAGAGTGGCGCAACTACTTAAAGAGCTTTTGGATATTCCTAT 175

160 AATCGAAATTTTGAAGGAGAAACCGGAAAGAAATATCTTGCAGTTTGGATGATCCAAC 219

176 GAAACATGCCAATTTGTTGAAGCTGATGATGAAATTTCAACGCTTGGAAATACCGTATCA 235

220 AAACATACAAATTTGCTGAAGCTGATGATGATGATGATGATGATGATGATGATGATGAT 279

236 CTTTGAACGGAGATTTGATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 295

280 TTTTGAACATGAGATTACGCAAGCATTCGACCATATTTATAGTGTATACGGTGTGATGATG 339

296 GAATGCTGACCGCTTTCTCTTATGCTTCTTATGCTTCTTATGCTTCTTATGCTTCTTATG 355

340 GAATGCTGCTGCTACTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 399

356 ATGTGATGCTTTTCAATAACTATAAAGCAAAATGAGCGTTTCAAGCAATTCGTTAGCTAA 415

400 ATGTGATGCTTTTCAATAACTATAAAGCAAAATGAGCGTTTCAAGCAATTCGTTAGCTAA 459

416 TGATGTTGAAGGTTTGTGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTT 475

460 GGTATTTGATGATGCTTGTGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGAT 519

476 TATATTAGAGATGCTTGTGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTT 535

520 CATTTAGATGAGGCTCTTGTGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGAT 579

536 TTTTCTTACAAACCGGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 595

580 ACTTGTGTTGCAACACAGCTCTCTAGACACATACATAGAGCACTAGAGCGGCTGTACA 639

596 GAAAGGTTGCAAGATAGAGGCGGCGAGTACATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 655

640 GAAAGGTTGCAAGATAGAGGCGGCGAGTACATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 699

656 TCATAACAGAGCTTTTACTTAACTTGTAGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTTAGTT 715

700 TCACAACAGGTCCTTACTTAAGCTTGCAGGCTTGGGGTTCAACCGGCTTCAATCCTTGCA 759

716 CAAGGAGAGCTCAGCCATGTTGCAATGTTGGAAGCTTTTCGATATCAAGAGAAACGC 775

760 TAAAGAGGCTTAGCCAACTTTCCAAATGTTGGAAGATTTGATGCTCCAAAGATCT 819

776 ACCTTGTTTAAGAGATAGATTTGTTGAATGCTACTTTTGGGAGCTAGGTTCTCAGGCTATGA 835

820 ACCTTACGTAAGAGATAGATTTGTTGAATGCTACTTTTGGGAGCTAGGTTCTCAGGCTATGA 879

836 GCCACAGTATTCGGGCTAGAGTTTCTTCAAAAAGCTGTTGCTGTTTATACTCTTAT 855

880 ACCTCAATATTTCTGTTGGAGATATTTCTTCAAAAAGCTGTTGCTGTTTATACTCTTAT 939

896 AGATGACACTTATGATGCTGATGCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 955

940 AGACGACAGTATGATGCTTATGATGCTTATGATGCTTATGATGCTTATGATGCTTATGATGCT 999

956 TGAAGGTTGCTCAATATCATGCTTTAGACACATTTCCAGATACATGAAACCGATACAA 1015

1000 TCAAGGTTGCTCAATATCATGCTTTAGACACATTTCCAGATACATGAAACCGATACAA 1059

```
Qy 1016 ATTATTCATGATACATACAGAAATGGAAGATTTCTTGCAAGGAGGGAAGACAGA 1075
Db 1060 GAGGCTCTTGATGTTTATGGAAGAAATGGAAGAAATCATAGAAAAGGATGGAAGACATA 1119
Qy 1076 TCTATTTAACTGCGGCAAGAAATTTGTGAAGAGTTTGTAGAAACCTGATGTTGAAGC 1135
Db 1120 TCAAGTTCACATGCAAAAGAGTCGATGATAGATTGGTTACAGATTATATGACCGAGC 1179
Qy 1136 AAAATGGCAATAGAGGACACATACCAACCACTGAAGAGCATGATCCAGTGTGAATCAT 1195
Db 1180 AAAATGGTTACATGAGGGTCATGTGCCAATTTGACGAGCATAACTCAGTTACAAACAT 1239
Qy 1196 TACTGGCGGTGCTTAACCTGTTACACAACTTTGTTATCTTGGCATG--AGTGATATATT 1252
Db 1240 AACTGGTGGCTATAAAATGCTTACAGCATCAAGCTTTGTGGCATGCGATGGTATATAGT 1299
Qy 1253 CACAAAAGAGTCTGTGCAATGGGCTGCTCTGCACCTCTCTTTTATAGTACTCAGGTAT 1312
Db 1300 TACCAAGAGCTTTCAATGGGTTCTCAACATCTCCACTTATAAAGCTTCATCTGA 1359
Qy 1313 ACTTGGTCGAGCCCTAAATGATCTCATGACCCCAAGGCCGAGCAAGAAAGAAACATAG 1372
Db 1360 CATTAGTAGGATTATGAATGATATCGTCCGTCAACAGGAGGAGCAACAAAGAAAGCATAT 1419
Qy 1373 TTCATGAGCCTTGAAGTTATATGAAGAAATATATGTCATGAGGAGTATGCCCAAC 1432
Db 1420 TGCACTAGTGTGGAATATGATGAAGAAATATATCTCGCGGAGGAGGAGCTCTATGA 1479
Qy 1433 CTGTGATTACAAAGAAAGTAGAAGATGTGTGGAAGATATAAACCAGAGAGTACCTCAAC 1492
Db 1480 TTTTCTCAAGAAAGAGTTGAAGTGCATGGAAGATATAACCGAGAAACATTACATG 1539
Qy 1493 TAAAAACATTCGAAGCCGTTATGATGGCTGTGATCTATTTGTGCGCATCTTCTTGAAGT 1552
Db 1540 TAAAGACATTCATATGGCTCTTAAAGATGCTCCGATCAACCTGGCAGCGGTAATGATAT 1599
Qy 1553 TCAATATCGAGGAAGGATACTTACAGCTATGGAGACGATACAAACATCTCATAA 1612
Db 1600 GCTATACAAAACCGTGATATTTTAAAAATGTGGACAAGAAATCCAGATTATATGAA 1659
Qy 1613 GTCTCTACTCGTTTATCTCTATGATATATGA 1643
Db 1660 ATCTGTTTCATTAATCTATGAGTGTGTA 1690

RESULT 9
LOCUS ATERPENE 1946 bp mRNA linear PLN 15-SEP-2000
DEFINITION Artemisia annua mRNA for epi-cedrol synthase.
ACCESSION AJ001539
VERSION AJ001539.1 GI:10183622
KEYWORDS epi-cedrol synthase.
SOURCE Artemisia annua (sweet wormwood)
ORGANISM Artemisia annua
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Asteroideae;
Anthemideae; Artemisia.
1
REFERENCE
AUTHORS Mercke,P., Crockett,J., Croteau,R. and Brodelius,P.E.
TITLE Cloning, expression, and characterization of epi-cedrol synthase, a
sequiterpene cyclase from Artemisia annua L
JOURNAL Arch. Biochem. Biophys. 369 (2), 213-222 (1999)
MEDLINE 99417501
PUBMED 10486140
REFERENCE
AUTHORS Brodelius,P.E.
TITLE Direct Submission
JOURNAL Submitted (15-SEP-1997) Brodelius P.E., Department of Plant
Biochemistry, Lund University, P.O.Box 117, Lund, S-22100, SWEDEN
FEATURES
Location/Qualifiers
1..1946
source
```

```
/organism="Artemisia annua"
/mol_type="mRNA"
/db_xref="taxon:35608"
/tissue_type="leaf"
7..10
34..1677
/codon_start=1
/product="epi-cedrol synthase"
/protein_id="CAC08805.1"
/db_xref="GI:10183623"
/db_xref="GOA:Q9LLR9"
/db_xref="SPREMBL:Q9LLR9"
/translation="MSLIVEDVIRPNFPEIWDGDFLAYDQDQEQVEQVVKDLKE
EVSELLTALNSPTQHTTELLKFIDAIRLGIAYVEEINOVFOHYMTAYGDKVTCGN
TSWFLMRQHGFEVSDSDISTYKDEGRFKESLEKDVHGLELYEAAMFYPVGGIL
HSLALFTRTCLDEIAKPSLNSAVSSOIREALTQPLHKRLPRLEALRYIPIYQOAS
HSTLLKXALGFNQLCSLHKLSELSISKWKSFDVANNLPYARNRPECYFWALAVY
FEPQSSRVFLSRFFSQTFLDDTYGTYTELEQFTEATQIRSIITCLDGLPESMK
LIFQMLVKIPFEEIEILSKGQKHVNYIKETLKAQVQSYTEARWAKEEYIPTIEEH
TKVSYLSIGYKLALVAGACWGDVIAADSPFEMVTNPPLVNACCLLCRTMDLGHKG
EQRKHVASITIECMKQFDASEQQAQYESLKNKVEDAKKEINREFMITCKDVNIHVNR
VLNFRSDVSLYIKNKHQFTHTGVEVINHIKSLFVDALIT"
1929
polyA_site
ORIGIN
Query Match 43.9%; Score 723.4; DB 8; Length 1946;
Best Local Similarity 65.4%; Pred. No. 7.7e-156;
Matches 1076; Conservative 0; Mismatches 566; Indels 3; Gaps 1;
```

```
Qy 2 CATGGCCTTACAGAGAAACCACTATTTCGCCCATTCCTCCCACTTTCTCCCAAGCATTTG 61
Db 33 CATGCTCTTATAGTAGAGGATGTCATACGCCCAATGCAAAATTTCTCTCTGAAATTTG 92
Qy 62 GGGAGATCAGTTCTCATCTATCAAAAGCAAGTAGAGCAAGGGGTGGACACATAGTGAA 121
Db 93 GGGAGATCAGTTCTTCGCTATGACAGGATGAGCAAGAGGAGTTGAAACAGTAAATCAA 152
Qy 122 TGATTTAAAAAGAGAGTGGCGCAACTACTAAAAAGAGCTTTGGATATTCTCTATGAAACA 181
Db 153 AGATTTGAAAGAGAGTGAAGCGGACCTATTGACAGCTCTAAATTTCTCCGACTCAGCA 212
Qy 182 TGCCAAATTTGTTGAAGCTGATGTAATGAAATTCACAGCCTTTGGAATACCGTATCATTGA 241
Db 213 TAGCGAGTGTGTTGAAATTTATTGACGCAATCAAGCGCTTGGTATTGGTATTATTGTA 272
Qy 242 ACGGAGATGATCATGATGCAATGTAATTTATGAAACATATGCTGATACCTGGAATGG 301
Db 273 AGAGGAGATCAACCAAGTATTTCACACATATATCTGCATATGCTGATGTAAGTGGACAGG 332
Qy 302 TGACCGCTCTTCCTTATGTTTCCGCTCTTATCGAAAGCAAGGATATTATGTTACATGTA 361
Db 333 TGGTAATACCTTCCTTTGTTTCCGCTCTTATCGACTCATGCGACATATGGAATTTTCGTTCAAGCGA 392
Qy 362 TGTTTTCAATAACTATAAAGACAAAATGGAAGCGTTCAAGCAATCGTTAGCTTAATGATGT 421
Db 393 TATTTTAGTACCTACAAAGACAAAGAGGAGCGTTTAAAGGAGTCTTTAGAAAAGGACGT 452
Qy 422 TGAAGTTTGGTTGAGTTGTACGAGCAACTCTATGAGGGTACCTCGGGGAGATTAATTT 481
Db 453 TCACGGGTACTTCAGTTGTATGAGGCGGCATATATGTTGTGCTGGGGAAGGGATACT 512
Qy 482 AGAGATGCTCTGTTTACACGATCTCGTCTTAGCATTTATGACAAAAGATGCTTTTTC 541
Db 513 AGATGATGCTCTGTTTACAGAACTGTCTTGTATGAAATAGCGAAAATTCCTAGTCT 572
Qy 542 TACAAACCCCGCTCTTTTACCGAAATAACAGGGGCACCTAAAGCAACCCCTTTGGAAAAG 601
Db 573 AAGTAATCTGCTGTTCTTCTCCCAATACGTGAAGCTCTAACGCAACCTTTACATAAAG 632
Qy 602 GTTGCCCAAGATAGAGGGCGGCGAGTACATCTCTTCTATCAACCAACAGATTCATATA 661
Db 633 ATTACCGAGACTAGAGGCGTTGCGCTACATACCTTTCTACCAACCAACAGCTTCTCACAG 692
```

RESULT 10					
AF157059					
LOCUS	AF157059	1985 bp	mRNA	linear	PLN 01-MAY-2001
DEFINITION	Artemisia annua 8-epicedrol synthase (Ecs1) mRNA, complete cds.				

ORIGIN	Query Match	43.9%;	Score 723.4;	DB 8;	Length 1985;
	Best Local Similarity	65.4%;	Pred. No. 7,7e-156;		
	Matches 1076;	Conservative	0;	Mismatches 566;	Indels 3; Gaps 1;
QY	2	CATGGCACTTACAGAGAAAACCTATTGCGCCCATTCGCAACTTTCTCTCCAGCAATTG	61		
Db	41	CATGTCCTTATAGTAGAGATGTCATAGCCCCCAATGCAAAATTTTCTCTGAAATTTG	100		
QY	62	GGGAGATCAGTTTCTCATCTATCAAAGCAAGTAGAGCAAGGGGTGGAAACAGATAGTGA	121		
Db	101	GGGAGATCAGTTTCTTTCGCTATGACCCAGGTAGCAAGAGGGAGTTGAAACAAGTAATCAA	160		
QY	122	TGATTTAAAAAAGAAGTGGCGCAACTACTAAAGAAGCTTTGGATATTCTCTATGAACA	181		
Db	161	AGATTTGAAAGAGAAGTGAAGAGCGAACTATTGACAGCTCTAAATCTCCGACTCAGCA	220		
QY	182	TGCCAATTTGTGAAGCTGATTCGATAAAATTCACGCTTGGATAACCGTATCACTTTGA	241		
Db	221	TACGGAGTTGTGAAATTTATTACGCGAATCGAACGCCCTTGGTATTGCGTATTATTTTGA	280		
QY	242	ACGGGAGATTGATCATGTCATTGCAATGTATTTATGAACAATATGGTGATAACTGGAATGG	301		
Db	281	AGAGGAGATCAACCAAGTATTCACACATGTATCTGCAATATGGTGATAAGTGACAGG	340		
QY	302	TGACCGCTCTTCCTTATGGTTCCGTTATGCGAAAGCAAGGATATTATGTTACATGTGA	361		
Db	341	TGCTAATACTTCCTTTTGGTTTCGACTCATGCGACACACATGGATTTTTCGTTTCAAGCGA	400		

QY 362 TGTGTTCAATTAATTAAGACAAATAATGAGCGTTCAGCAATCGTTAGCTAATGATGT 421
 Db 401 TATTTTATGACCTACAAAGACAAAGAGGAGCGTTTAAAGGAGCTTTTAAAGAGGAGCGT 460
 QY 422 TGAAGGTTGCTGAGTTGACGAGCAACTTCTATGAGGGTACCTGGGAGATATATTT 481
 Db 461 TCACGGGTTACTTGGTGTATGAGGGGCGATATATGTTTGGCTGGGAGGAGTACT 520
 QY 482 AGAAGATGCTTGTGTTTACACGATCTCGTCTTAGCATTTATGACAAAGATGCTTTTTC 541
 Db 521 AGATGATGCTTGTGTTTACAGAACTTGTCTGTATGAAATAGGAAATCTTAGTCT 580
 QY 542 TACAAACCCCGCTCTTTTACCGAAATACAAAGGCACTAAAGCAACCCCTTTGGAAG 601
 Db 581 AAGTAATCTGCTTGTCTTCCCAATACGTAAGCTCTAAGCAACCTTTACATATAAG 640
 QY 602 GTTCCCAAGATAGAGCGGGCGAGTACATCTCTTCTATCAACCAAGATTTCTCATAA 661
 Db 641 ATTACCGAGACTAGGCGGTTGCGCTACATACCTTTCTACCAACCAAGCTTTCTCAG 700
 QY 662 CAAGACTTTACTTTAACTTGTCTAAGTTAGAGTCTCAATTTGCTTCACTTGCACAAAG 721
 Db 701 TGAGACATTTGTTAACTGCTGCAATTTAGGTTTCAACCACTTCACTTGCACAAAG 760
 QY 722 AGAGCTCAGCCATGCTGCAAAATGCTGAAAGCTTTTCGATATCAAGAAAGACGACCTTG 781
 Db 761 AGAGCTTAGCATATTTCCAGTGTGGAAGTTTCGATGTCGAAACAATCTTACCGTA 820
 QY 782 TTTAAGAGATAGATTTGTTGAATGCTTCTTTGGGACTAGTTCAGGCTATGAGCCACA 841
 Db 821 TGCAAGAAATAGACCCGTTGAATGCTTCTTTGGCATTAGCAGTGTACTTCGAGCCCTCA 880
 QY 842 GTATTCGCGCTAGAGTTTCTTCAAAAAGCTTGTCTGCTTAACTTATAGATGA 901
 Db 881 ATATCTGAATCAAGAGCTTCTTATCAAGATTTTCTTCAATACAAATCTTCTGATGA 940
 QY 902 CACTTATGATGCTGATGCTTACTTATGAAGAACTTAAAGTCTTTACTGAGCTGTTGAAG 961
 Db 941 CACTTATGATGCTGATGCTTACTTACGAAGAACTTGAACAACTTACTGAGCAATTTCAA 1000
 QY 962 GTGCTCAATACATGCTTTGAGCGCTTCTTCTGAAAGCATGAGCTTATTTTCAAAATGCT 1021
 Db 1001 GTGCTCAATACATGCTTTGAGCGCTTCTTCTGAAAGCATGAGCTTATTTTCAAAATGCT 1060
 QY 1022 CATGATACATACACGAAATGGAAGATTTCTTCAAGAGGAGGAGCAAGACATCTATT 1081
 Db 1061 AGTGAAGATTTTGAAGAAATAGAGGAATATTTGCGAAGGATGGAAGACAGCATCATGT 1120
 QY 1082 TAAGTGGGCAAGAAATTTGTGAAGAGTCTTGTGAGAAACCTGATGTTGAGCAAAATG 1141
 Db 1121 CAATCATCAAGAAACACTGGAAGAGGCTGTCGAAGCTACATGACAGAGCAAGATG 1180
 QY 1142 GGCAAAATGAGGACACATACCAACACTGGAAGAGCATGATCCAGTTGTAATCACTTACTGG 1201
 Db 1181 GGCAAAATGAGGATATATACCAACGATTTGAGGAGCAGCAACAAAGTGTCTATATTTAGCAT 1240
 QY 1202 CGGTGCTAACTGCTTACAACTGTTGATCTTGGCATGAGTATATTTACAAAGAA 1261
 Db 1241 CGGTGCTAACTGCTTACAACTGTTGATCTTGGCATGAGTATATTTACAAAGAA 1300
 QY 1262 GTCTGTCAATGAGGCTGCTCTGCACTCTCTTTTATAGATCTCAGGTATATCTTGGTGG 1321
 Db 1301 CTGATTTGAATGGGTTTACCAATCTCTCTTCTGATGCTTGTGCTACTTTGTTAG 1360
 QY 1322 ACGCTAAATGATCTCATGACCAAGGCGGAGCAAGAAAGAAACATAGTTTCATCGAG 1381
 Db 1361 AACCATGATGATCTTGGTCTCCCAATAGGCGGAGCAAGATAGAAAGCATGTTGCTCTCC 1420
 QY 1382 CTTGAAAGTTATATGAGGATATATGTTCAATGAGGAGTATGCCAAACCTTGTATTTA 1441
 Db 1421 TATTGAATGCTACATGAAGCAATTTGATGCTAGGAGCAACAGGCTTACGATCACTCAA 1480

QY 1442 CNAAGAGTAGAGATGCTGGAAGATATATACCGAGAGTAC--CTCACAACTAAAAA 1498
 Db 1481 CAAGAAAGTTGAAGATGTCATGGAAGAAATAATAGAGAGTTTATGATCATGCTTAAGGA 1540
 QY 1499 CATTCCAAGCGCTTATTCATGCTGCTGATCTATTTGTCAGTCTTCTTGAAGTTCAATA 1558
 Db 1541 TCTTATATTCATGATGCTATGAGGCTCTTAAATTTTTCAGGTTGCGTGGATGCTCTTTA 1600
 QY 1559 TCGAGAAAGATATACTTCACAGTATGGAGACGAATACAAATCTCATAAAGTCTCT 1618
 Db 1601 TAAGAAATAGGATCACTTTTACACATGCTGGAGTAGAGGTGATTAATCATATCAATCTCT 1660
 QY 1619 ACTCGTTTATCTTATGATGATGA 1643
 Db 1661 TTTCTGTTGATGCTCATTTACATGA 1685

RESULT 11
 AF327527

LOCUS
 DEFINITION

ACCESSION
 AF327527

VERSION
 AF327527.1

KEYWORDS
 GI:13195433

ORGANISM
 Artemisia annua (sweet wormwood)

REFERENCE
 Liu, Y., Ye, H.C. and Li, G.F.

TITLE
 Cloning of sesquiterpene cyclase gene from Artemisia annua

REFERENCE
 Liu, Y., Ye, H.C. and Li, G.F.

TITLE
 Direct Submission

JOURNAL
 Submitted (08-DEC-2000) Department of Cell and Gene Engineering,
 Institute of Botany, Chinese Academy of Sciences, Xiangshan,
 Nanxincun 20, Beijing 100093, P.R. China

FEATURES
 Location/Qualifiers

1. .4392

/organism="Artemisia annua"

/mol_type="genomic DNA"

/strain="001"

/db_xref="taxon:35608"

/tissue_type="leaf"

/country="China: Sichuan"

<13. .33978

/genes="AMS1"

join(13. .99,214. .484,2183. .2564,2741. .2959,3088. .3226,
 3324. .3572,3685. .3978)

/product="amorpho-4,11-diene synthase"

join(13. .99,214. .484,2183. .2564,2741. .2959,3088. .3226,
 3324. .3572,3685. .3978)

/genes="AMS1"

/notes="sesquiterpene cyclase"

/codon_start=1

/product="amorpho-4,11-diene synthase"

/protein_id="AAK15697.1"

/db_xref="GI:13195433"

/translation="MSLITEKPIRPIANFPPIWMDQFLIYKEQVQGVQVNDLKK
 EVQLKEALDIPMKHANLLKIDEIOLGTPHYFEQIDHALQCIYETGNDWGD
 SSWFLMRKQGVYVTCDFNNYKDKGAFKOSLANDVEGLLELYEATSMRVPQIML
 EDALGTRSELSTMTKDAFSTNPALFTEIOLAKQPLWKRLPRTEAOYIIFYQOQDS
 HNTLKLALEFNLOSLHKBELSHVCKWKAEDIKNAPCLDRIVCEVFWGLGSG
 FEPQGRARVFTTAVATLIDDTDAYGTIEELKFTPEAVERNSITCLDILPYMK
 PIVKPMDTYEMEEFLKRGDTDFNCGEFVKEFVFNLMVEAKWANEHGIPTTEH
 DPVILITGGANLLTTCYLGMSDITFKESVEWAVSAPLFRYSGLGRNLDMTHKA
 EQRKSSSLESYMKYNNVEVAOTLIYKEVEDVWKNINREVLTKNIPRLIMAV
 IYLCQLEVOYAGKDNFTRNGDEYKHLIKSLLYVPMST"

gene

mRNA

CDS

ORIGIN

Query Match 43.9%; Score 723.2; DB 8; Length 4392;
Best Local Similarity 71.1%; Pred. No. 8.1e-156;
Matches 1281; Conservative 0; Mismatches 8; Indels 513; Gaps 4;

QY	360	GATGTTTCAATACTATAAAGACAAAATGAGCGTTCACGCAATCGTTAGCTAATGAT	419
DB	2182	GATGTTTCAATACTATAAAGACAAAATGAGCGTTCACGCAATCGTTAGCTAATGAT	2241
QY	420	GTTGAAGGTTTGGTTGAGTTGACGAAGCAACTCTATGAGGGTACCTGGGGAGATTATA	479
DB	2242	GTTGAAGGTTTGGTTGAGTTGACGAAGCAACTCTATGAGGGTACCTGGGGAGATTATG	2301
QY	480	TTAGAAGAGTCTCTGGTTTACAGCATCTGCTTAGCAATATGACAAAAGATGCTTTT	539
DB	2302	TTAGAAGAGTCTCTGGTTTACAGCATCTGCTTAGCAATATGACAAAAGATGCTTTT	2361
QY	540	TCTCAAAACCCCGCTCTTTTACCAGAAATACAAAGGCACTAAAGCAACCCCTTTGGAAA	599
DB	2362	TCTCAAAACCCCGCTCTTTTACCAGAAATACAAAGGCACTAAAGCAACCCCTTTGGAAA	2421
QY	600	AGGTTGCAAGAATAGAGCGCGCAGTACATTCCTTTCTATCAACAACAAGATTCTCAT	659
DB	2422	AGGTTGCAAGAATAGAGCGCGCAGTACATTCCTTTCTATCAACAACAAGATTCTCAT	2481
QY	660	AACAAGACTTTTAACTTGTCTAAGTAGAGTTCAATTTGCTTCAGTCATGCAACAG	719
DB	2482	AACAAGACTTTTAACTTGTCTAAGTAGAGTTCAATTTGCTTCAGTCATGCAACAG	2541
QY	720	GAAGAGCTCAGCCATGTGTGCAA	742
DB	2542	GAAGAGCTCAGCCATGTGTGCAAGTATATACATATAGCCTCAACCAATCTTAAATTA	2601
QY	743	-----	742
DB	2602	GTTTTATTCTTTCAAATTTTATATTTTATTTCTATGCTCTTTTGGTTTCTTATTCT	2661
QY	743	-----	742
DB	2662	CATTATTACTTCATCAGTATCTACAAATTAAGGAAAGAAATAAGCTAACAAAAACATA	2721
QY	743	-----ATGGTGGAAAGCTTTTGGATATCAAGAAAGACGCAACCTTGT	783
DB	2722	ATTATGATGTTACATAGGTGGTGGAAAGCTTTTGGATATCAAGAAAGACGCAACCTTGT	2781
QY	784	TAAAGATAGAAATTTGAAATGCTACTTTTGGGACTAGGTTCAAGCTATGAGCCACAGT	843
DB	2782	TAAAGATAGAAATTTGAAATGCTACTTTTGGGACTAGGTTCAAGCTATGAGCCACAGT	2841
QY	844	ATTCGGGCTAGAGTTTCTTCACAAAAGCTTTGCTGTTTAACTCTTATAGATGACA	903
DB	2842	ATTCGGGCTAGAGTTTCTTCACAAAAGCTTTGCTGTTTAACTCTTATAGATGACA	2901
QY	904	CTTATGATGCGTATGCTACTTATGAAGAACTTAAGATCTTTTACTGAAGCTGTTGAA	959
DB	2902	CTTATGATGCGTATGCTACTTATGAAGAACTTAAGATCTTTTACTGAAGCTGTTGAAAGT	2961
QY	960	-----	959
DB	2962	ATTTGTTAAACATGCTACATGTCACACCTTACACACATACATATATTAACATA	3021
QY	960	-----	959
DB	3022	GAACCTTCAATCTATTAACCAACCATGTATATAAATCAAACTAAATACGCAAGCTCAATTA	3081
QY	960	-----AGGTGGTCAATTTACATGCTTAGACACACTTCCAGAAATACATGAAACCGATATACAA	1015
DB	3082	CGACAGGTGGTCAATTTACATGCTTAGACACACTTCCAGAAATACATGAAACCGATATACAA	3141
QY	1016	ATTATTCATGGATACATACACAGAAATGGAAGAAATTTCTTCAAAAGGAGGGAAGAACAGA	1075
DB	3142	ATTATTCATGGATACATACACAGAAATGGAAGAAATTTCTTCAAAAGGAGGGAAGAACAGA	3201

QY	1076	TCTATTTAACTGGCGCAAGAAATTT	1100
DB	3202	TCTATTTAACTGGCGCAAGAAATTTGTAAGTATCAACACCAATCTTTGATATATTTTAAAA	3261
QY	1101	-----	1100
DB	3262	AAAAAATCCTGATGAAGATAATATATGTAAGACACGCTGACTGTATGCAATGTGTCATGCAC	3321
QY	1101	--GTGAAAGAGTTTGTAGAAACCTGATGTTGAAGCAAAATGGGCAAAATGAGGACACA	1158
DB	3322	AGGTGAAAGAGTTTGTAGAAACCTGATGTTGAAGCAAAATGGGCAAAATGAGGACACA	3381
QY	1159	TACCAACCACTGGAAGAGCATGACCTGCTTAATCATTTACTGCGGGTGTCAACCTGCTTA	1218
DB	3382	TACCAACCACTGGAAGAGCATGACCTGCTTAATCATTTACTGCGGGTGTCAACCTGCTTA	3441
QY	1219	CAACAACTTGTATCTTGGCATGAGTGATATATTTCAAAAAGAGTCTGTGGAATGGGCTG	1278
DB	3442	CAACAACTTGTATCTTGGCATGAGTGATATATTTCAAAAAGAGTCTGTGGAATGGGCTG	3501
QY	1279	TCTCTGACCTCTCTTTTGTAGATACATGAGTATATCTTGTGTCGAGCCCTAAATGATCTCA	1338
DB	3502	TCTCTGACCTCTCTTTTGTAGATACATGAGTATATCTTGTGTCGAGCCCTAAATGATCTCA	3561
QY	1339	TGACCCACA	1347
DB	3562	TGACCCACAAGGTACACCCCTTTTCTAGTAGATATATCTATCTTTAACTATTTCTGTTTC	3621
QY	1348	-----	1347
DB	3622	TAAACTCTAGTTTATGCGCGAGCCTAAATAACCAATTTATTTTGTAAATATAATTTAT	3681
QY	1348	-AGGCGAGCAAGAAAGAAAACATAGTTTCATGAGCCTTGAAGTTATATGAGGAAATAT	1406
DB	3682	CAGGCGAGCAAGAAAGAAAACATAGTTTCATGAGCCTTGAAGTTATATGAGGAAATAT	3741
QY	1407	AATGTCAATGAGGAGTATGCCAAACCTTGATTTTACAAGGAAGTAGAAGATGTGTGAAA	1466
DB	3742	AATGTCAATGAGGAGTATGCCAAACCTTGATTTTACAAGGAAGTAGAAGATGTGTGAAA	3801
QY	1467	GATATAAACCGAGAGTACCTCACAACTTAAACCAATTTCCAGGCGCTTATTCATGCTGTG	1526
DB	3802	GATATAAACCGAGAGTACCTCACAACTTAAACCAATTTCCAGGCGCTTATTCATGCTGTG	3861
QY	1527	ATCTATTGTCAGCTTTCTTGAAGTTCAATATGACGAAAGGATTAATTCACACGATG	1586
DB	3862	ATCTATTGTCAGCTTTCTTGAAGTTCAATATGACGAAAGGATTAATTCACACGATG	3921
QY	1587	GGAGACGAATACAAACATCTCATAAAGTCTCTACTCGTTTATCCTATGAGTATAGAGA	1646
DB	3922	GGAGACGAATACAAACATCTCATAAAGTCTCTACTCGTTTATCCTATGAGTATAGACTA	3981
QY	1647	TC 1648	
DB	3982	CC 3983	

RESULT 12

AAN249561
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

AAN249561
Artemisia annua mRNA for putative sesquiterpene cyclase.
AJ249561
AJ249561.1 GI:5921216
sesquiterpene cyclase.
Artemisia annua (sweet wormwood)
Artemisia annua
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;
asterids; campanulids; Asterales; Asteraceae; Asteroideae;
Anthemideae; Artemisia.
1
Van den Beekhout, E., Van Geldre, E., De Pauw, I. and Inze, D.

Qy	542	TACAAACCCGGCTCTTTTATCCGAAATACAAACGGGCACTAAAGCAACCCCTTTGGAAG	601
Db	573	AAGTAACCTGCTATTCTTCTCCAAATACGTGAAGCTCTAAGCAACCTTTTACATAAAG	632
Qy	602	GTTCGAAGTATAGAGCGGGAGTACATTCCTTCTATCAACAACAAGATTCCTATAA	661
Db	633	ATTACCGAGACTAGAGCGGTGGCTACATACCTTTTACCAACAACAAGTTCTCACAG	692

Qy	662	CAAGACTTTACTTTAAACTGCTAAGTTAGAGTTCAATTGCTTTCAGTCAATTCACACAGGA	721
Db	693	TGAGACATTTGTTAAACCTGGCCAAATTAGGGTTCAACCAACTTCAATCACTTCCACAGAA	752
Qy	722	AGAGCTCAGCCATGTTGTCAAATGGTGGAAAAGCTTTCGATATATCAAGAAAGAACGCACCTTG	781
Db	753	AGAGCTTAGCATTAATTTCCCAAGTGGTGAAAAGTTTCGATGTTGCAAAACAATCTACCGTA	812

753	AGAGCTTAGCATAAATTTCCAAAGTGGTGAAGAAAGTTTCGATGTTGCAAAACAATCAACGTA	812
Db		
782	TTTAAAGAGATAGAAATGTTGAATGCTACTTTTGGGGACCTAGGTTCAAGGCTATCAGGCACA	841
Qy		
813	TGCNAGAAATAGACCCGTTGAAATGCTACTTTTGGGCATTAGCAGTGCTACTTCGAGCCTCA	872
Db		
842	GTATTCCTCCGGCTAGAGTTTCTTCCAAAAGCTGTTGCTGTTATTAACCTTTATAGATGA	901
Qy		
873	ATATCTCGAATCAAGAGTCCTCTTATCAAGATTTTTTCAATCAACAACATTTCTTGTGATGA	932
Db		
902	CACATTATGATCGGTATGGTACTTATGAAGAACTTAAAGATCTTTTACTGAAGCTGTGAAAG	961
Qy		
933	CACATTATGATCGCTACGGTACTTACGAAGAACTTGAACAATTTTACTGAAGCAATTCAAAG	992
Db		
962	GTGTCCTAATTACATGCTTTAGACACACTTCCAGAATACATGAACCGATATACAAATATT	1021
Qy		
993	GTGTCCTAATAACATGCTTTGAGCGTCTTCTGAAGCATGAAGCTTATATCAAAATGCT	1052
Db		

[illegible]

Qy	1082	TAAC	TGCGGCAAGAAATTG	TGAAAGAGTTGGT	TAAGAACCTCGATG	GTGGTAAAGGAACTG	TTG	1172
Db	1113	CAACT	TACATCAAGAAACACT	CAAGAGGCTGT	CCAAAGCTACAT	GCACAGAAGCAAGATG		1172
Qy	1142	GGCAAT	GAGGACACATACCA	CCACCACTGAAGAGCAT	CCAGTTCTTAATCAT	TACTACTGG		1201
Db	1173	GGCAAAA	GAAAGAGTATATACCA	ACGATGAGGAGACAC	AAAAGTGCATATAT	TAGCAT		1232
Qy	1202	CGGTCTAAC	CTGCTTACAA	CAACTTGTATCTTTGG	CATGAGTGAATAT	TTCACAAAAGA		1261
Db	1233	CGGCTAT	ATAACTTCGCTAG	TAGCAGGTTTGCT	TGCATGGGTGAT	GTAAATTGCAGATGA		1292
Qy	1262	GTCTCTCGAAT	GGGTGTCTCTGC	ACTCTCTTTTAGAT	ACTCAGGTATAC	TGTGGTGC		1321
Db	1293	CTCATTTGAAT	GGGTTTACC	CAATCTCTCTTGT	CAATGCTTGTGTC	TACTTTGTAG		1352
Qy	1322	ACGCTTAAT	GATCTCATGCC	ACCAAGGCGAGCA	AGAAAGAAACATAG	TTTCATCGAG		1381
Db	1353	AACCAT	TGGATGATCTTTGG	TCCCATAGGGCGAGCA	GATAGAAGCATGTT	GCCTCCAC		1412
Qy	1382	CCCTGAAAG	TTATATGAAGAA	TATATATGTCAAT	TGAGGAGTATGCC	AAAACCTTGATTTA		1441
Db	1413	TATTGAA	TGCTACATGA	GCAAATTTGATGCTAG	CGAGCAACAGCTAC	GAACTCACTCAA		1472
Qy	1442	CAAGGAAG	TAGAAGATGTGTG	GAAGAAGATATAAAC	CGGAGAGTAC	---CTCACAACTATAAAA		1498
Db	1473	CAAGAAG	TTGAGATGCTATG	GAAGAAGAAATAAT	TAGAGAGTTCATG	ATCACGTGTAAAGA		1532
Qy	1499	CATTCCA	AGGCCGTTATTGAT	GGCTGTGATCTATT	TGTGCCAGTTTCTT	TGAAGTTCAATA		1558
Db	1533	TGTTAA	TATATCATGTAG	TATGAGGTCCTTAATTTT	CAGTTCGGTGGATG	TCTCTTTA		1592
Qy	1559	TGCAGGA	AAAGGATAACTT	CTCATCGAT	TGGGAGACGAATA	CAAAAACATCTCAT	AAAGTCTCT	1618
Db	1593	TAAGAAT	AAGGATCACTTTT	TACATGTTGGAG	TAGAGGTGATAAAT	TCATATCAAAATCTCT		1652

```

QY 1619 ACTCGTTTATCTCTATGATATATGA 1643
Db 1653 TTTCGTTGATGCCATCATATGA 1677

RESULT 13
AF304444 AF304444 1886 bp mRNA linear PLN 23-FEB-2001
LOCUS Artemisia annua sesquiterpene cyclase mRNA, complete cds.
DEFINITION Artemisia annua sesquiterpene cyclase mRNA, complete cds.
ACCESSION AF304444
VERSION AF304444.2 GI:131113759
KEYWORDS
SOURCE Artemisia annua (sweet wormwood)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Asteroideae;
Anthemideae; Artemisia.
1 (bases 1 to 1886)
Liu Y., Ye H.C. and Li, G.F.
Cloning, characterization and E.coli expression of a cDNA encoding
sesquiterpene cyclase from Artemisia annua
2 (bases 1 to 1886)
Liu Y., Ye H.C. and Li, G.F.
Direct Submission
Submitted (11-SEP-2000) Department of Cell and Gene Engineering,
Institute of Botany, Chinese Academy of Sciences, No. 20 Nanxin
Chun, Xiangshan, Beijing 100093, P.R. China
3 (bases 1 to 1886)
Liu Y., Ye H.C. and Li, G.F.
Direct Submission
Submitted (23-FEB-2001) Department of Cell and Gene Engineering,
Institute of Botany, Chinese Academy of Sciences, No. 20 Nanxin
Chun, Xiangshan, Beijing 100093, P.R. China
Sequence update by submitter
On Feb 23, 2001 this sequence version replaced gi:10945664.
FEATURES
location/Qualifiers
1..1886
/organism="Artemisia annua"
/mol_type="mRNA"
/strain="001"
/db_xref="taxon:35608"
/country="China: Sichuan"
42..1763
/codon_start=1
/product="sesquiterpene cyclase"
/protein_id="AAG24640.2"
/db_xref="GI:13113760"
/translation="MSTFLVCNAPSFPSPYGLPSSCDNSNMQEIYRNTVKPPSIWG
DPLITHEPKSLTIKQHVVEEIRKLLNTAFSEPKOHTLKLHLIDSVQRLGVSY
HPEQIEEALQHYATHGDQWIGKNLSTSQWFRILRQGFNVSSGIFENHMDKQK
FKEHLDDVQGMGLYEAAYMSVEGEKELDDALEFTKHLGNIAKOPSONASLRKIE
QALNRPLRMRPLREALHYIPKYQOASHDELTLLNKLIDYNNLQSIHKREIIEICKW
WDLDPSNKLPHVRDLRLVIFYFWLGVYFEPQHYRSMFLTKTSMLLVLDDYDNYG
TYEELKIIFYEVRWLSCLDSLPYMKLIYRELINHYQWESSEVEKESYQLFYAK
EMVKNCRLLVEAKKEGVYPTLEHMSVSCVTYAYVNIANSIVGRNNWVTEESF
KWAKFPPLVNLATCLLIRLMDIDAGHEEOERDHSVSSIECYRNETGASEEDAVKFLS
KQVEDAWKVINKECLEPTEIPMSLTIPPLSLARVSDILYKTNNGYNHAGEVISYKS
LLVHPLV"

CDS
Query Match 35.7%; Score 588; DB 8; Length 1886;
Best Local Similarity 61.9%; Pred. No. 1.le-124;
Matches 1011; Conservative 0; Mismatches 600; Indels 23; Gaps 4;

QY 28 TTGCGCCCATTCGCACTTTCCTCCAGCATTTGGGGAGATCAGTTTCTCATCTATCA-- 85
Db 136 TCGGCAACTGTATAATTTCTCTAGCATATGGGGAGATCAGTTTCTTACATACCATG 195
QY 86 -----AAAGCAAGTAGCAAGGGGTGGAAACAGATAGTGAATGATTTAAAAAAGAAAGTG 140
Db 196 AACCGAAAAAAGTTTAAACCATTTGAGAAACAAACATGTTGAAGAGGTTGAAGAAATAAG-G 254

141 CGGCAACTACTAAAAAGAGCTTTTGGATATTCCTATGAAACATGCCCAATTTGTTGAAGCTG 200
255 AAAAAGCTATTGAAACACAGCTTTTAGTGAGCCAAAAGCAGCATACGAAAGCTAATACATCTT 314
201 ATTGATGAATTCAGCGCTTGGATACGGTATCACTTTGAAGCGGAGATTGATCATGCA 260
315 ATAGATTGAGTTCAAGCGCTTGGAGTATCTTATCATCTTTTGAACAAGAAATTTAGGAAGCC 374
261 TTGCAATGATTTATGAACAATATCGTGTACTACTGGAATGG-----TGACCGC 308
375 TTGCAGCATGCTATGCTACACATGCTGATCAGTGGATTGGCAAGCAACCTGAAAGT 434
309 TCTTCTTATGTTCCGCTCTTATGCGAAAGCAGAGATATATGTTACATGTTGATGTTTTC 368
435 ACATCACAATGTTCCGAATCCTTAAGACACAAGGTTTAAATGTTTTCATCTGGAATCTTC 494
369 AATAACTATAAAGACAAAAATGGAGCGTTCAGCAATCGTTAGCTAATGATGTTGAAGGT 428
495 GAGAACCATATGATGACAGGAAATTTAAGGAACACTTGAGTGACGATGTTCAAGGC 554
429 TTGCTTGAAGTTGACGAAGCAACTTCTATGAGGGTACCTGGGGAGATTTATATTAGAAGAT 488
555 ATGCTTGGTTTGTATGAAGCAGCATATATGAGTGTGAAGCGGAAAAAGAAATAGATGAT 614
489 GCTCTTGGTTTACACGATCTGCTTACGATATATGACAAAAGATGCTTTTCTACAAAC 548
615 GCCCTCGAATTCACATAAAGCTCACCCTTGAAACATAGCAAGGATCC---TTCTCAGAAT 671
549 CCCGCTCTTTTACCGAAATACAAAGCGCCTTAAAGCAACCCCTTTGAAAAAGGTTGCCA 608
672 GCTTCTTTAAGAACCAAAATAGAACAGCATTAACAGACCACTTCGGAAGAGATGCCA 731
609 AGAATAGAGGGCGCGCATGATCTCTTCTATCAACAAACAGATTTCTCATACAGACT 668
732 AGGCTAGAGGCACGTGCTATACATACCTAAATACCAACAAAGAGCTTCTCATGACGAGAC 791
669 TTACTTAACTTGTAGCTTAGAGTTCAATTTGCTTCTAGTCAATGACACAGGAGAGCTC 728
792 TTGTTAAATCTTGCAAGTGTGATTTATACATGCTTCAATCAATACAAAAAGGAGAT 851
729 AGCCATGTGTGCAAAATGTTGAAAGCTTTTCGATATCAAGAAAGAACCCACTTGTGTTAAGA 788
852 AGTGAATCTGCAAAATGTTGAGAGATTTGGATTTCTCAACAAACACTACCGCATGTTCGA 911
789 GATAGAAATGTTGAATGCTACTTTTGGGGAATGAGTTGAGGCTATGAGCCACAGTATTCC 848
912 GATAGATGTTGAAATATATTTTGGGTATTTGGGGTCTATTTTGAACCTCAACATTTAT 971
849 CGGCTAGAGTTTCTTCCAAAGCTGTTGCTGTATTAACCTTTATAGATGACACTTAT 908
972 CGTTCAGGATGTTTCTCACCAAAACGAGATATGTTGTTAATTTGTTCTGATGACATAT 1031
909 GATGCTATGCTACTTATGAAGAACTTAAGATCTTTTACTGAAGCTGTTGAAAGGTTGCTCA 968
1032 GATAATATGTTACTTACGAGAACTTAAGATCTTTTACGAGCTGTTGAGAGATGTTCC 1091
969 ATTCATGCTTAGACACACTTCCAGAAATACATGAACCGATATACAAATTTATTCATGGAT 1028
1092 TTAAGCTGTTGGACTCACTTCCAGATACATGAAATTAATATATCGAAGAACTTATAAAT 1151
1029 ACATACACAGAAATGGAAGATTTCTTCCAAAGGGAGGAGAACAGATCTATTATACTGC 1088
1152 CATTACCAGGAAATGGAAGATCAGTTGAAAGGAGGAAATCATATCAATTTATTCTAT 1211
1089 GGCAAGAAATTTGTCAAGAGTTTCTTAGAAACCTGATGTTGGTGAACAAAAATGGGCAAT 1148
1212 GCTAAGAGATGTTAAGAAATTTGTGAAACCTTGTAGTGAAGCAAAAATGTTTAAAA 1271
1149 GAGGACACATACCAACCACTGAAGAGCATGATCAGTTGTTAATCATTTACTGGCGGTGCT 1208
1272 GAAGGATACGTACCAACCACTTGAGGAACACATGTCGGTTTCTATGTTTACTTATGCCTAT 1331

```

```
QY 1209 AACCTGCTTACAACTGTTATCTTGGCATGAGTATATATACAAAAAGAGTCTGTC 1268
Db 1332 GCTGTCAATGATCGCAACTCTTATGTTGGACGGGAAATATGTTTACAGAGGATCCCTTT 1391
QY 1269 GAATGGGCTGCTCTGACCTCTCTTTTATAGATACCTCAGGTATATCTGGTCAGCCCTA 1328
Db 1392 AAGTGGGCTTGCCTTCTCTCTCTTCTTAAATGCTACATGTTTGAATTTAAGGCTTATG 1451
QY 1329 AATGATCTCATGACCCACAGCCGAGCAAGAAAGAAACATAGTTTCATCGAGCCCTTGA 1388
Db 1452 GATGATATCGGGCCACGAGGAGAAACAAAGAGAGACCATGTTGTTTCAAGTATTGAA 1511
QY 1389 AGTTATATGAAGGAATATAATGTCATGAGGAGTATGCCAAACCTTGATTTACAAGGAA 1448
Db 1512 TGCTACAGAAATGAACTGGAGCTCGAGAGGAGTGGTAAAGTTCTTTCAAAACA 1571
QY 1449 GTAGAGATGTTGGAAGATATATAACCGAGAGTACCTCAACATTAACAACTTCAAGG 1508
Db 1572 GTTACAGATGATGGAAGTTATAAATAAGAGTGTCTTAAGCCCTACAGAAATTCCTCATG 1631
QY 1509 CGTTATGATGGCTGTGATCTATTGTCAGGTTTCTTGAAGTTCAATATCGAGGAAG 1568
Db 1632 TCTCTAACTATACCTCCACTCAGCCTTGCTCGTGTGAGTGATATCCTTTATAAAACGAAT 1691
QY 1569 GATAACTTCAACGTATGGGAGAGCAATACAAACATCTCATAAAGTCTCTACTCGTTTAT 1628
Db 1692 AATGTTACATATCAGGAGGGAAGATGATCAGTTACATTAAGTTCGCTCTGTTTCTAT 1751
QY 1629 CCTATGATGATG 1642
Db 1752 CCTTTAGTTTAATG 1765

RESULT 14
AAN271792 1843 bp mRNA linear PLN 04-OCT-2000
LOCUS Artemisia annua mRNA for putative sesquiterpene cyclase (cASC125
DEFINITION gene).
ACCESSION AJ271792
VERSION AJ271792.1 GI:10696987
KEYWORDS cASC125 gene; sesquiterpene cyclase.
SOURCE Artemisia annua (sweet wormwood)
ORGANISM Artemisia annua
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Asteroideae;
Anthemideae; Artemisia.
REFERENCE 1
AUTHORS Van Geldre, E., De Pauw, I., Inze, D., Van Montagu, M. and Van den
Beckhout, E.
TITLE Cloning and molecular analysis of two new sesquiterpene cyclases
from Artemisia annua L
JOURNAL Plant Sci. 158 (1-2), 163-171 (2000)
PUBMED 10996256
REFERENCE 2
AUTHORS Van Geldre, E.G.L.
TITLE Direct Submission
JOURNAL Submitted (31-JAN-2000) Van Geldre E.G.L., Laboratory for
Pharmaceutical Biotechnology, University of Ghent, Harelbekestraat
72, 9000 Gent, BELGIUM
FEATURES
source
1..1843
/organism="Artemisia annua"
/mol_type="mRNA"
/db_xref="taxon:35608"
/tissue_type="leaves"
1..1843
/gene="cASC125"
61..1794
/gene="cASC125"
/codon_start=1
/evidence="experimental"
/product="putative sesquiterpene cyclase"
gene
CDS
```

```
/protein_id="CAC12731.1"
/db_xref="GI:10696988"
/db_xref="GOA:Q9FYX7"
/db_xref="SPTREMBL:Q9FYX7"
/translation="MSTLPISSVSFSSTSSPLVDDKVSTKPDVIRHTMTNPNASWGD
QFLTDEPDLVMKKQLVEELKEVKELITIKSNPEPMQHVKLIELIDAVORLGIAY
HFEETEEALQIHVITYGQWVDKENLQISLWFLRLQQQFNYSVGVFKDFMDEKIGK
HFEELCNDAGILALVEAAMRVEDTILDNALFETKVHLDI IAKDPSCDSLRQIKH
QALKPLRLARIEALHVMPIYQOETSHNEDLLKAKLDFSVLQSMHKKLSLICKW
WKLQDLQNLFPVDRVRVEGYFWLISYIPEPQHARTFMFLMKTCKMLVLDLDPNYG
TYEELIEFTQAVERSISCLMDPEYNKLIYQELVNLHVEHVESVSVMTGTGLMARSYVGRGDIVTE
GWQKELGSOITLIVETMAKRGVHAQPLEEYMSVMTGTGLMARSYVGRGDIVTE
DTFKWYSVPPIIKASCVIIRLMDDIVSHKEQERGHVASSIECYSEKESAE
YTSRKVEDAWKVINRESLRPTAVFPFLMLPAINLARMCEVLVSNDGFTHAEGDMKSY
MKSFFVHPWV"
ORIGIN
Query Match 33.9%; Score 559; DB 8; Length 1843;
Best Local Similarity 61.6%; Pred. No. 5.2e-118;
Matches 1025; Conservative 0; Mismatches 605; Indels 33; Gaps 7;
QY 12 ACAGAAAGAAACCTATTGCCCCCATTCGCAACTTTCTCCAGCATTTGGGGAGATCAG 71
Db 136 ACAGAACCGAGTGTATCGCCATACCATGATTTCAATGCTAGCATATGGGAGATCAA 195
QY 72 TTTCTCATCTATCAAAAGCAAGTAGCAAGGGTGG---AACAGATAGTGAATGATTTA 128
Db 196 TTTCTTACATATGATGAGCCAGAGATTGGTTATGAAGAAACAACTGTTGTTGAAGAGCTG 255
QY 129 AAAAAAGAAAGTCGGCAACTTACTATAAAG-----AAGCTTTGGATATTCCTATGAAACAT 182
Db 256 AAAGAGGAAGTCAAAAAGAGAGTAAATAACGATCAAGAGTTCAAAATGAACCGATGAGCAT 315
QY 183 GCCAATTTGTTGAAGCTGATTGATGAATAATTCACGCCCTTGGAAATACCGTATCAGTTTGA 242
Db 316 GTCAAGTTGATAGAACTTATTGATGAGTCCACGGCTTGCGCATGAGCTATCATTTTGA 375
QY 243 CGGGAGATTGATCATGATTCGATTCATTTATTTGAAACATATGTTGATAACTGG----- 296
Db 376 GAGGAGATTGAAGAGCGTTGCAACATATCCATGTTACATATGCGGAGCAGTGGTTCAT 435
QY 297 -----AATGGTGACCGCTCTCTCTTATGTTCCGTCTTATGCGAAAGCAAGGATATTAT 350
Db 436 AAGAAACCTACAAAGCAATTTCTTTGTTCCGACTCTCTACGACAAAGGCTTCAAC 495
QY 351 GTTACATGATGATTTTCAATAATCTATAAGACAAATAATGAGGCGTTCAAGCAATCGTTA 410
Db 496 GTCTCATCAGGAGTATTCAAAGACTTTTATGACGAGAAAGGAAAAATTTAAGGAATCCTTA 555
QY 411 GCTAATGATGTTGAAGGTTTGGCTTGTAGTTGTACGAGCAACTCTATGAGGCTACCTGG 470
Db 556 TGTAATGATGCTCAGGGAATCTTCTCTGTATGAAGCGGCATTTATGAGGTTGAAGAT 615
QY 471 GAGATTATATTAGAAGATGCTCTTGGTTTACAGCATCTCGTCTTAGCATTATGACAAA 530
Db 616 GAAACATACTAGATATGCGCTCGAGTTTCAATAAGTTTCACTTACATCATAGCCAA 675
QY 531 GATGCTTTTCTACAAACCCCGCTCTTTTACCGAAATACAAACGGGCACTAAGCAACCC 590
Db 676 GATCCTTCTTGT---GACTCTTCCCTAAGAACCCAAATACACCAAGCGCTTAAGCAGCCA 732
QY 591 CTTTGAAGAGGTTGCCAGAAATAGAGCGCGGCGAGTACATTCCTTCTATCAACAA 650
Db 733 CTTAGGAGAGGCTGGCAAGGATGAGCGGTGTGATTCATTCATTCATTCACCAAGAA 792
QY 651 GATTCTCATACAAAGACTTTTACTTAACTTGTAAAGTTAGAGTTCAATTTGCTTCAGTCA 710
Db 793 ACATCCCAACATGAGGACTTACTGAGCTTGCATAATTAGATTTTAGCGTTCAGTCG 852
QY 711 TTGCAACAGGAAGAGCTCAGCCATGTGTCAAAATGTTGGTGAAGAGCTTTCGATATCAGAG 770
Db 853 ATGCACAAAAGGAAGAACTTAGCCACATTTGCAAAATGGTGAAGAGGATTTGAGCCTCCAAAT 912
```

771 AACGACCTGTTAAGAGATAGAAATGTTGAATGCTACTTTTGGGACATAGTTCAGGC 830
913 AAGCTTCCTTATGTTTCGAGACAGGGTGTGGAGGCTATTTTGGATACATATCCACTAT 972
831 TATGAGCACACAGTATTCGGGGCTAGAGTTTCTTTCACAAAGCTGTTGCTGTATACT 890
973 TACGAGCCTCAACATGCTCGAACAAAGATGTTCTTAATGAAACATGCTAGTGTAGTC 1032
891 CTTATAGATGACACTTATGATCGTATGCTATGCTTATGAGAACTTTAAGATCTTTACTGAA 950
1033 GTTTTAGATGATACATTTTGATAATATGCTACTTATGAGAACTCGAAATCTTTACACA 1092
951 GCTGTTGAAGGTGCTCAATACATGCTTAGACACATCTCCAGATACATAGAACGATA 1010
1093 GCTGTTGAAGATGCTCAATAAGCTGCTCGGATATGCTTCCAGAAATACATAGAACTAA 1152
1011 TACAAATTTATTCATGATACATACACAGAAATGGAAGATTTCTT-GCAAAGGAGGAG 1069
1153 TATCAGAGCTTGTAATCTTCACGTGGAATGAGGAATCACTTGGAAAGAGGAGAA 1212
1070 AACGATCTATTTAACTCGGCAAGAAATTTGTGAAGA-----GTTTGTAGAAACCT 1123
1213 AACATATCAATTTCACTATGTCAAGGAGATGCGAAAGAGTGTGGTTTCGCAAAATTACT 1272
1124 GATGTTGAAGCAAAATGGGCAATGAGGACACATACC--AACCACTGAAGAGCATGAT 1181
1273 TTGTAGAAACCAAGATGGCTAAAAGAGGGGTACATGCGCAACCTCTTTGAGGAGTACATG 1332
1182 CCAGTTGTAATCATTTACTTACCTGCTTACCAACCTGTTTATCTTGGCATG 1241
1333 TCTGTGCAATGGTACGTGTTACCTATGCTTGTATGATAGCAGATCTTATGTGGCAGG 1392
1242 AGTGATATATTTACAAAAGAGTCTCGAATGGGCTCTCTGCACTCTCTCTTTTAGA 1301
1393 GGTGATATCGTCAACGAGATACCTTTTAAATGGTGTCTCTGATCTCTCTTATATAAAA 1452
1302 TACTCAGTATATCTTGTGCGAGCCTCAATGATCTCATGCCACCAAGCCGAGCAAGAA 1361
1453 GCTTCATGTTGTAATGTAAGCTTTATGATGATATTTGTCAGCCACCAAGGAGCAAGAG 1512
1362 AGAAACATAGTTTCAATGAGCTTGAAGTATATGAGGAATATAATGTCATGAGGAG 1421
1513 AGAGCCATGTTGTTCAAGCATGCAATGCTATTTCAAGGAAGTGGTGCATCAGAGGAG 1572
1422 TATGCCCAAACTTGATTTACAGGAAGTAGAAGATGTTGTGGAAGATATAAACCGAGAG 1481
1573 GAAGGCTGTAATATATCTCAAGAAAAGTTGAAGATGATGAAAGCTTATAAACCGAGAG 1632
1482 TACCTCAACAATAAAACATTCAGAGCCGTTTATGATGCTGTGATCTATTGTGCCAG 1541
1633 TCGCTCAGGCTTACAGCTGTCCCATTTTCTCTACTATGCTGCAATCAACCTTGCACGG 1692
1542 TTTCTTGAAGTTCAATATGCGAGGAAGGATACTTCAACAGCTATGGGAGACGAATACAAA 1601
1693 ATGTGGAAGTCTTATATAGCTCAACGATGCTTCACCTCATGCTGAGGAGACATGAAA 1752
1602 CATCTCAATAAGTCTTACTCTGTTATCTTATCCTATGAGTATAGAG 1644
1753 AGTTACATGAATCTGTTCTTCTGTTACCCCTATGTTGTTGCTAAG 1795

RESULT 15
BD227439
LOCUS BD227439 538 bp DNA linear PAT 17-JUL-2003
DEFINITION Transgenic amorpho-4,11-diene synthesis.
ACCESSION BD227439
VERSION BD227439.1 GI:33037209
KEYWORDS JP 2002523101-A/1.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 538)
AUTHORS Wallaart,T.E. and Bouwmeester,H.J.

Transgenic amorpho-4,11-diene synthesis
Patent: JP 2002523101-A 1 30-JUL-2002;
GENOLIPP BIOTECHNOLOGY BV
OS Artemisia annua L.
EN JP 2002523101-A/1
PD 30-JUL-2002 JP 2000567711
PF 27-AUG-1998 EP 98202854.0
PI THORVALD BELCO WALLAART,HENDRIK JAN BOUWMEESTER PC
C12N15/09,C12N15/09,A01H5/00,C12N1/19,C12N1/21,C12N5/10,C12N9/ PC
88,C12P5/00,
PC C12P17/18,C12N9/04,C12N1/19,C12R1/84),(C12N1/19,C12R1:645),
PC (C12N9/04,C12R1:91),(C12N9/04,C12R1:865),(C12N9/04,C12R1:84),
PC (C12N9/04,C12R1:19),(C12N9/04,C12R1:645),(C12N9/88,C12R1:91),
PC (C12N9/88,C12R1:84),(C12N9/88,C12R1:84),(C12N9/88,C12R1:865),
PC (C12N9/88,C12R1:645),(C12N15/09,C12R1:91),(C12P5/00,C12R1:19),
PC (C12P5/00,C12R1:84),(C12P5/00,C12R1:91),(C12P5/00,C12R1:865),
PC (C12P5/00,C12R1:645),(C12P17/18,C12R1:645),(C12P17/18,C12R1:19) PC
PC
PC (C12P17/18,C12R1:91),(C12P17/18,C12R1:84),(C12P17/18,C12R1:865) PC
PC C12N15/00,C12N5/00,(C12N15/00,C12R1:91)
CC PCR product amplified on cDNA as template with primers CC
CC starting and ending
CC at nucleic acid positions (1..23 and 517..538) FH Key
FT CDS Location/Qualifiers
1..538
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 31.0%; Score 510.6; DB 6; Length 538;
Best Local Similarity 97.4%; Pred. No. 7.9e-107;
Matches 519; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 381 GACAAAATCGAGCGTTCAGCAATCGTTAGCTAATGATGTTGAAGTTCCTTGAGTTG 440
DB 1 GATGAGATCGGAAATTTAAGGAATCGTTAGCTAATGATGTTGAAGTTCCTTGAGTTG 60
QY 441 TAGAAGCAACTCTTATGAGGTACCTGGGGAGATATATTAGAGATGCTCTTGTTTT 500
DB 61 TAGAAGCAACTCTTATGAGGTACCTGGGGAGATATATTAGAGATGCTCTTGTTTT 120
QY 501 ACAGATCTCGTCTTAGCATTATGACAAAAGATGCTTTTCTACAAACCCGCTCTTTTT 560
DB 121 ACAGATCTCGTCTTAGCATTATGACAAAAGATGCTTTTCTACAAACCCGCTCTTTTT 180
QY 561 ACCGAATACACGGGCACTAAAGCAACCCCTTTGGAAAAGTTGCCAAGATAGAGCG 620
DB 181 ACCGAATACACGGGCACTAAAGCAACCCCTTTGGAAAAGTTGCCAAGATAGAGCG 240
QY 621 GCGAGTACATTCCTTCTATCAACAACAGATTCATATACAGACTTTTAACTT 580
DB 241 GCGAGTACATTCCTTCTATCAACAACAGATTCATATACAGACTTTTAACTT 300
QY 681 GCTAAGTTAGAGTTCAATTTGCTTCACTTGCACAGGAAGAGCTCAGCCATGTGTC 740
DB 301 GCTAAGTTAGAGTTCAATTTGCTTCACTTGCACAGGAAGAGCTCAGCCATGTGTC 360
QY 741 AAATGTGGAAGCTTTTCGATATCAAGAAGAACCCCTTCTTTTAAAGATAGAAATGTT 800
DB 361 AAATGTGGAAGCTTTTCGATATCAAGAAGAACCCCTTCTTTTAAAGATAGAAATGTT 420
QY 801 GAATGCTACTTTTGGGCACTAGTTTGGCTATGAGCCACAGTATTCGGGCTAGAGTT 860
DB 421 GAATGCTACTTTTGGGCACTAGTTTGGCTATGAGCCACAGTATTCGGGCTAGAGTT 480
QY 861 TTCTTCACAAAGCTGTTGCTGTTTAACTCTTATAGATGACACTTATGATGC 913

TITLE
JOURNAL
COMMENT

Db 481 TTCTTCACAAAAGCTGTGCTGTATATAACTCTTATAGACGACACCTTCGACGC 533

Search completed: June 7, 2004, 11:23:28
Job time : 6517 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2004, 08:48:09 ; Search time 735 Seconds
(without alignments)
9530.991 Million cell updates/sec

Title: US-09-763-822A-13

Perfect score: 1649

Sequence: 1 ccatggcattacagaagaa.....ctatgagtatataggatgcc 1649

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_29Jan04:*

1: Geneseq1980s:*

2: Geneseq1990s:*

3: Geneseq2000s:*

4: Geneseq2001as:*

5: Geneseq2001bs:*

6: Geneseq2002s:*

7: Geneseq2003as:*

8: Geneseq2003bs:*

9: Geneseq2003cs:*

10: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1649	100.0	1649	3	Aaz92754 Artemisia
2	1640.6	99.5	2112	3	Aaz92760 Artemisia
3	1637.4	99.3	2076	3	Aaz92763 Artemisia
4	1637.4	99.3	2076	3	Aaz92757 Artemisia
5	1637.4	99.3	2076	3	Aaz92757 Artemisia
6	1637.4	99.3	2076	3	Aaz92757 Artemisia
7	1637.4	99.3	2076	3	Aaz92757 Artemisia
8	1637.4	99.3	2076	3	Aaz92757 Artemisia
9	1637.4	99.3	2076	3	Aaz92757 Artemisia
10	1637.4	99.3	2076	3	Aaz92757 Artemisia
11	1637.4	99.3	2076	3	Aaz92757 Artemisia
12	1637.4	99.3	2076	3	Aaz92757 Artemisia
13	1637.4	99.3	2076	3	Aaz92757 Artemisia
14	1637.4	99.3	2076	3	Aaz92757 Artemisia
15	1637.4	99.3	2076	3	Aaz92757 Artemisia
16	1637.4	99.3	2076	3	Aaz92757 Artemisia
17	1637.4	99.3	2076	3	Aaz92757 Artemisia
18	1637.4	99.3	2076	3	Aaz92757 Artemisia
19	1637.4	99.3	2076	3	Aaz92757 Artemisia
20	1637.4	99.3	2076	3	Aaz92757 Artemisia
21	1637.4	99.3	2076	3	Aaz92757 Artemisia
22	1637.4	99.3	2076	3	Aaz92757 Artemisia
23	1637.4	99.3	2076	3	Aaz92757 Artemisia

24	225.6	13.7	1650	2	AAX35575	Aax35575 Computer
25	225.6	13.7	1650	2	AAX35574	Aax35574 Computer
26	225.6	13.7	1650	2	AAX35576	Aax35576 Computer
27	225.6	13.7	1650	2	AAX35578	Aax35578 Computer
28	225.6	13.7	1959	3	AAX38926	Aax38926 Mint E-b-
29	225.6	13.7	1960	2	AAX35570	Aax35570 DNA encod
30	224.8	13.6	1650	2	AAX35573	Aax35573 Computer
31	224.8	13.6	1650	2	AAX35577	Aax35577 Computer
32	223.2	13.5	1910	7	ADA71275	Ada71275 Rice Gene
33	219.4	13.3	1976	3	AAC38691	Aac38691 Arabidops
34	217.2	13.2	1728	7	ADA68331	Ada68331 Arabidops
35	205.6	12.5	1524	7	ADA69380	Ada69380 Rice gene
36	204	12.4	1575	7	ADA70943	Ada70943 Rice gene
37	199.2	12.1	1631	9	ADE82680	Ade82680 Terpenoid
38	199.2	12.1	1874	9	ADE82682	Ade82682 Terpenoid
39	199.2	12.1	1874	9	ADE82678	Ade82678 Terpenoid
40	198.2	12.0	1665	9	ADE82745	Ade82745 Terpenoid
41	196	11.9	1665	9	ADE82747	Ade82747 Terpenoid
42	193.4	11.7	1894	9	ADE82684	Ade82684 Terpenoid
43	191.6	11.6	1560	7	ADA70344	Ada70344 Rice Gene
44	183.6	11.1	1806	7	ADA69387	Ada69387 Rice Gene
45	161.4	9.8	1878	3	AAC49451	Aac49451 Arabidops

ALIGNMENTS

RESULT 1
AAZ92754
ID AAZ92754 standard; cDNA; 1649 BP.
XX AC AAZ92754;
XX DT 19-JUN-2000 (first entry)
XX DE Artemisia annua amorpho-4,11-diene synthase cDNA.
XX KW Amorpho-4,11-diene synthase; amorphadiene synthase; FPP;
KW farnesyl pyrophosphate; cyclisation; quinghao; sweet wormwood;
KW annual wormwood; sweet annie; artemisin precursor; antimalarial;
KW sesquiterpene; flavouring; fragrance; ss.
XX OS Artemisia annua.
XX FH Key Location/Qualifiers
FT CDS 3..1643
FT /tag= a
FT /product= "Artemisia annua amorpho-4,11-diene synthase"

authentic
inverted

CC amorpha-4,11-diene. Amorphadiene is a precursor of artemisin, a
CC sesquiterpene lactone endoperoxide produced by *A. annua* which is a
CC promising antimalarial drug candidate. Amorphadiene synthase is a branch
CC point enzyme in the artemisin biosynthetic pathway, and the cyclisation
CC of FPP is therefore likely to be the rate-limiting step. Artemisin is
CC present in very low concentrations in *A. annua*, making it expensive for
CC use as an antimalarial drug, and attempts at organic synthesis have been
CC unsuccessful. Sesquiterpenes such as amorphadiene are also useful as
CC flavour and fragrance compounds in the food and perfume industries.
CC Terpenes also play a role in plant-insect interactions, such as the
CC attraction or repulsion of insects by plants. In addition,
CC dihydroarcteanic acid, an intermediate in the metabolic route from
CC amorphadiene to artemisin in *A. annua*, can be used as an antioxidant.
CC DNA encoding amorphadiene synthase may be used to generate transgenic
CC plants able to synthesise amorphadiene synthase. Preferably, the plants
CC used for amorphadiene production are plants which naturally produce
CC sesquiterpenes, as these plants already have the basic sesquiterpene
CC synthetic pathways and storage compartments. *A. annua* may be transformed
CC with amorphadiene synthase expression constructs, enabling more
CC amorphadiene to be produced and hence increasing artemisin production. It
CC may then be economically feasible to extract artemisin from such plants
CC for use as an antimalarial
XX
SQ Sequence 1649 BP; 537 A; 299 C; 350 G; 463 T; 0 U; 0 Other;
Query Match 100.0%; Score 1649; DB 3; Length 1649;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1649; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCATGGCACTTACAGAGAAAACCTATTGCCCCATTGCCAACTTCTCCTCCAAACATTT 60
DB 1 CCATGGCACTTACAGAGAAAACCTATTGCCCCATTGCCAACTTCTCCTCCAAACATTT 60
QY 61 GGGGAGATCAGTTTCTCATCTATCAAAAGCAAGTAGAGCAAGGGGTGAAACAGATAGTGA 120
DB 61 GGGGAGATCAGTTTCTCATCTATCAAAAGCAAGTAGAGCAAGGGGTGAAACAGATAGTGA 120
QY 121 ATGATTTAAAAAAGAGTGGCGCACTACTAAAGAGAGCTTTGGATTTCTCTATGAAC 180
DB 121 ATGATTTAAAAAAGAGTGGCGCACTACTAAAGAGAGCTTTGGATTTCTCTATGAAC 180
QY 181 ATGCCAATTTGTTGAAGCTGATGATGAAATTCAGCGCTTGGATACCGTATCACTTTG 240
DB 181 ATGCCAATTTGTTGAAGCTGATGATGAAATTCAGCGCTTGGATACCGTATCACTTTG 240
QY 241 AACGGGAGATGTATGATGCAATGCAATGATTTATGAACAATATGGTGATTAACGGAAATG 300
DB 241 AACGGGAGATGTATGATGCAATGCAATGATTTATGAACAATATGGTGATTAACGGAAATG 300
QY 301 GTGACCGCTCTTCTTATGTTCCGTTTCCGTTTATGCGAAGCAAGGATATATGTTACATGTG 360
DB 301 GTGACCGCTCTTCTTATGTTCCGTTTCCGTTTATGCGAAGCAAGGATATATGTTACATGTG 360
QY 361 ATGTTTTCATTAATACTAATAAGACAAAATGAGCGGTTCAAGCAATCGTTAGCTAATGATG 420
DB 361 ATGTTTTCATTAATACTAATAAGACAAAATGAGCGGTTCAAGCAATCGTTAGCTAATGATG 420
QY 421 TTGAAGTTTCTTGTAGTTGTGAGAACCACTTCTATGAGGGTACCTGGGAGATTATAT 480
DB 421 TTGAAGTTTCTTGTAGTTGTGAGAACCACTTCTATGAGGGTACCTGGGAGATTATAT 480
QY 481 TAGAAGATGCTCTTGTGTTTACAGATCTCGTCTTAGCATTTATGACAAAAGATGCTTTT 540
DB 481 TAGAAGATGCTCTTGTGTTTACAGATCTCGTCTTAGCATTTATGACAAAAGATGCTTTT 540
QY 541 CTACAAACCCCGCTCTTTTACCGAAATACAAACGGGCACTAAAGCAACCCCTTTGGAAA 600
DB 541 CTACAAACCCCGCTCTTTTACCGAAATACAAACGGGCACTAAAGCAACCCCTTTGGAAA 600
QY 601 GGTGGCAAGAAATAGAGCGCGGCGAGTACATCTCTTTCTATCAACAAAGATTTCTCATA 660
DB 601 GGTGGCAAGAAATAGAGCGCGGCGAGTACATCTCTTTCTATCAACAAAGATTTCTCATA 660

QY 561 ACAAGCTTTTACTTAAACTGCTTAAGTTAGAGTTCAATTTGCTTCACTTCACTGCAAGG 720
DB 561 ACAAGCTTTTACTTAAACTGCTTAAGTTAGAGTTCAATTTGCTTCACTTCACTGCAAGG 720
QY 721 AAGAGCTCAGCCATGTGTCGAATGTGGAAGCTTTCGATATCAAGAAAGAACGACCTT 780
DB 721 AAGAGCTCAGCCATGTGTCGAATGTGGAAGCTTTCGATATCAAGAAAGAACGACCTT 780
QY 781 GTTTAAGAGATGAATTTGTAATGCTACTTTTGGGAGTCTAGGTTTCAAGGCTATGAGCCAC 840
DB 781 GTTTAAGAGATGAATTTGTAATGCTACTTTTGGGAGTCTAGGTTTCAAGGCTATGAGCCAC 840
QY 841 AGTATTCGGGCTAGAGTTTCTTCAAAAAGCTGTTGCTGTTTAACTTCTTATAGATG 900
DB 841 AGTATTCGGGCTAGAGTTTCTTCAAAAAGCTGTTGCTGTTTAACTTCTTATAGATG 900
QY 901 ACACCTTATGATCGGTATGCTTCTTATGAAGAACTTAAAGATCTTCTTAAAGCTTCTGAAA 960
DB 901 ACACCTTATGATCGGTATGCTTCTTATGAAGAACTTAAAGATCTTCTTAAAGCTTCTGAAA 960
QY 961 GGTGTCATTAATCATGCTTTAGACACACTTCCAGAAATACATGAAACCGATATACAAATAT 1020
DB 961 GGTGTCATTAATCATGCTTTAGACACACTTCCAGAAATACATGAAACCGATATACAAATAT 1020
QY 1021 TCATGATACATACACAGAAATGGAAGAAATTTCTTCRAAGAGGGAGGAACAGATCTAT 1080
DB 1021 TCATGATACATACACAGAAATGGAAGAAATTTCTTCRAAGAGGGAGGAACAGATCTAT 1080
QY 1081 TTAATCTCGGCAAGAAATTTGTAAGAGTCTTGTAGAAACCTGATGGTTGAAGCAAAAT 1140
DB 1081 TTAATCTCGGCAAGAAATTTGTAAGAGTCTTGTAGAAACCTGATGGTTGAAGCAAAAT 1140
QY 1141 GGGCAAAATGAGGAGACACATACCAACCACTGAAGAGCATGATCCAGTTGTAATCATTTACTG 1200
DB 1141 GGGCAAAATGAGGAGACACATACCAACCACTGAAGAGCATGATCCAGTTGTAATCATTTACTG 1200
QY 1201 GGGTGTCAACCTGCTTCAACCACTGTTTATCTTGGCATGAGTGATATATTCACAAAG 1260
DB 1201 GGGTGTCAACCTGCTTCAACCACTGTTTATCTTGGCATGAGTGATATATTCACAAAG 1260
QY 1261 AGTCTGTGGAATGGGCTGCTCTGCAACCTCTCTTTTAGATATCTCAGGTATATCTTGGTC 1320
DB 1261 AGTCTGTGGAATGGGCTGCTCTGCAACCTCTCTTTTAGATATCTCAGGTATATCTTGGTC 1320
QY 1321 GACGCTAAATGATCTCTGATGACCCCAAGCCGAGCAAGAAAGAAACATAGTTTCATCGA 1380
DB 1321 GACGCTAAATGATCTCTGATGACCCCAAGCCGAGCAAGAAAGAAACATAGTTTCATCGA 1380
QY 1381 GCTTTGAAAGTTTATGAGGAAATATATGTCATGAGGAGTATGCCAAACCTTGAATTT 1440
DB 1381 GCTTTGAAAGTTTATGAGGAAATATATGTCATGAGGAGTATGCCAAACCTTGAATTT 1440
QY 1441 ACAAGCAAGTAGAAGATGTGGAAGATATAAAACGAGAGTACCTCACAACTAAAAACA 1500
DB 1441 ACAAGCAAGTAGAAGATGTGGAAGATATAAAACGAGAGTACCTCACAACTAAAAACA 1500
QY 1501 TTCCAAGCCGCTTATGATGCTGATCTATTTGTCAGGTTTCTTGAAGTTCAATATG 1560
DB 1501 TTCCAAGCCGCTTATGATGCTGATCTATTTGTCAGGTTTCTTGAAGTTCAATATG 1560
QY 1561 CAGGAAAGGATTAACCTTCAACGATGAGGAGACGATACAAACATCTCTATAAAGTCTCTAC 1620
DB 1561 CAGGAAAGGATTAACCTTCAACGATGAGGAGACGATACAAACATCTCTATAAAGTCTCTAC 1620
QY 1621 TCGTTTATCTCTATGATATATGAGGATCC 1649
DB 1621 TCGTTTATCTCTATGATATATGAGGATCC 1649

RESULT 2
AAZ92760
ID AAZ92760 standard; cdna; 2112 BP.
XX

amorphadiene to be produced and hence increasing artemisin production. It may then be economically feasible to extract artemisin from such plants for use as an antimalarial. The present sequence represents a positive cDNA clone encoding amorphadiene synthase identified in an *A. annua* CDNA library using a 538 bp amorphadiene synthase probe (AZ9257) in an exemplification of the present invention

Sequence 2112 BP; 689 A; 376 C; 423 G; 624 T; 0 U; 0 Other;

Query Match	99.5%;	Score 1640.6;	DB 3;	Length 2112;
Best Local Similarity	99.8%;	Pred. No. 0;		

```
Matches 1643; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

2 CATGGCACATTCAGAGGAAATACCTATTGGCCCAATGGCCCAACTTTCTCTCAAGACATTTC 61

24 CAGTCACCTTACGCGATGGTTTATTCCGCGCATTTGCCGTAATTCCTCCTGCGGCACTTGC
62 GCGAGATCATGTTTTCTCATCTATCTAAAAAGCAAGTTAGAGCAGGGGGTGGAAACAATAGTGAA

84 GGAGATCAGTTTCTCATTATCAAAGCAAGTACAGCAAGGGTGGACAGATAGTAA 143

122 TGATTTAAAAAAGAAGTGGGCAACTACTAAAGAAGCTTTGGATATTTCCTATGAACA 181

144 TGATTTAAAAAGAAAGTCGGCACTACTAAAAGACCTTTGGATATTCTATGAACA 203

182 TGCCAATTTGTTGAAGCTGATTGATGAAATTCACGCCTTGGAAAPACCGTATCACTTTGA 241

204 TGCCAAATTGTTGAAGCTGATGTGTAATTCACGCCCTTGGAAATACCGTATCACATTGA 263

242 ACGGAGATTGATCATGCATTGCAATGATTTATGAACAATATGGTGATTAACCTGGAATGG 301

284 ACGGGAGATTTGATCATGCAATGGCAATGATTTATGATATCATATGGTGATATACATGGAAATGG 323

324 TGACCGCTCTTCTTATGGTTCCGTCCTTATGCGAAGCAAGGATATTATTGTACATGTGA 383

362 TGTGTTCAATAACTATATAAGACAAAAATGGAGCGCTCAAGCAATCGTTAGCTAATGATGT 421

384 TGTTTCAATAACTATAAGACAAAAATGGAGCGTTCAAGCAATCGTTAGCTAATGATGT 443

422 TGAAGGTTTCTTCAGTTCTACGAAGCAACTTCTATGAGGGTACCTGGGAGATTATATT 481

444 TGAAGTTTGCTTGAGTTGTACGAAGCAACTTCTATGAGGTACTCTGGGAGATTATATT 503

482 AGAAGATGCTCTGGGTTTACACGATCTCGTCTTAGCATTTATGACAAAAGATGCTTTTTC 541

542 TACAGACCCGCGCTCTTTTATACGAGATACAAACGCGCACTAAGCAACCCCTTTTGGAAAAG 601
504 AGAAGATGCTCTTTGTTTTTACACGATCTGCTTAGCATATGACAAAAGATGCTTTTTC 563

564 TACAAACCCCGCTCTTTTACCGAAATCAACGGGCACTAAAGCAACCCCTTTGGAAAAG 623

602 GTTGCCAGAATAGAGCGCGGAGTACATTCTTTCTATCAACAACAGATTCTCATAA 661

624 GTTGCCAGAAATAGAGCGCGCAGTACATTCCTTTCTATCAACAACAAGATTCATATA 693

662 CAAGACITTTACTTTAAACITTGCTAAGTTAGAGTTCAATTTTGCTTCAGTCAATTCACACAGGA 721

684 CAAGACTTTACTTTAAACCTTGCTAAGTTAGAGTTCAATTTTGCTTCAGTCAATTGCACAAGGA 743

722 AGAGCTCAGCCATGTTGTGCAAAATGGTGGAAAGCTTTTCGATATCAAGAAGACGCACCTTG 781

744 AGAGCTCAGCCATGTGTGCAAAATGGTGGAAAGCTTTCGATATCAAGAAGAAACGCACCTTG 803

762 TTTAAGAGATAGATATGTTGAAATGCTACTTTTGGGACATAGGTTTCAAGGCTATAGGCCACA GTT
804 TTTAAGAGATAGATATGTTGAAATGCTACTTTTGGGACATAGGTTTCAAGGCTATAGGCCACA GTT
863

842 GTATTCCGGGGCTAGAGTTTCTTCACAAAGCTGTTGCTTATTAACCTTTATAGATGA 901

864 GTATTCCGGGGTAGAGTTTCTTCACAAAAGCTGTTGCTGTTAATCTTATAGATGA 923

QY 902 CACTTATGATCGTATGCTACTTATGAGAACTTAAGACTTTTACTGAAGCTTTGAAAG 961
Db 924 CACTTATGATCGTATGCTACTTATGAGAACTTAAGACTTTTACTGAAGCTTTGAAAG 983
QY 962 GTGGTCAATTACATCTTATGACACACACTTCCAGAAATACATGAACCGATATACAAATTAT 1021
Db 984 GTGGTCAATTACATCTTATGACACACACTTCCAGAAATACATGAACCGATATACAAATTAT 1043
QY 1022 CATGGATACATACACAGAAATGGAAGAAATTTCTTGCAAAAGGAGGGAAGAACAGATCTATT 1081
Db 1044 CATGGATACATACACAGAAATGGAAGAAATTTCTTGCAAAAGGAGGGAAGAACAGATCTATT 1103
QY 1082 TAACCTGCGCAAAAGAAATTTGTAAGAGGTTTGTAGAAAACCTGATGGTTGAAGCAAAATG 1141
Db 1104 TAACCTGCGCAAAAGAAATTTGTAAGAGGTTTGTAGAAAACCTGATGGTTGAAGCAAAATG 1163
QY 1142 GGCAAAATGAGGACACATACCAACCACTGAAAGGATGATCCAGTTGTATCATCTACTCG 1201
Db 1164 GGCAAAATGAGGACACATACCAACCACTGAAAGGATGATCCAGTTGTATCATCTACTCG 1223
QY 1202 CGGTGCTAACCTGCTTACAAACCTGTTATCTTGGCATGAGTGATATTCACAAAAGA 1261
Db 1224 CGGTGCTAACCTGCTTACAAACCTGTTATCTTGGCATGAGTGATATTCACAAAAGA 1283
QY 1262 GTCTGTGGAATGGGCTGTCTCTGACCTCTCTTTTATGATCTACAGGTATCTTGGTCTG 1321
Db 1284 GTCTGTGGAATGGGCTGTCTCTGACCTCTCTTTTATGATCTACAGGTATCTTGGTCTG 1343
QY 1322 AGCCTTAATGATCTCATGACCCACAGGCGGAGCAAGAAAGAAACATAGTTCATCGAG 1381
Db 1344 AGCCTTAATGATCTCATGACCCACAGGCGGAGCAAGAAAGAAACATAGTTCATCGAG 1403
QY 1382 CTTTGAAGGTTATATGAAGGAATATAATGCTCAATGAGGAGTATGCCAAACCTTGATTTA 1441
Db 1404 CTTTGAAGGTTATATGAAGGAATATAATGCTCAATGAGGAGTATGCCAAACCTTGATTTA 1463
QY 1442 CAAGGAAGTAGAAGTGTGTGGAAGATATAAACCAGAGAGTACCTCACAACTAAAAACAT 1501
Db 1464 CAAGGAAGTAGAAGTGTGTGGAAGATATAAACCAGAGAGTACCTCACAACTAAAAACAT 1523
QY 1502 TCCAGGCGGTTATGATGGCTGTGATCTATTGTGCGCAGTTCTTGAAGTTCAATATGC 1561
Db 1524 TCCAGGCGGTTATGATGGCTGTGATCTATTGTGCGCAGTTCTTGAAGTTCAATATGC 1583
QY 1562 AGGAAGGATTAACCTCACACGATGCGGAGACGAATACAAACATCTCTAAAGTCTCTACT 1621
Db 1584 AGGAAGGATTAACCTCACACGATGCGGAGACGAATACAAACATCTCTAAAGTCTCTACT 1643
QY 1622 CGTTTATCCTATGAGTATATGAGGATC 1648
Db 1644 CGTTTATCCTATGAGTATATGAGTATC 1670

RESULT 3
AAZ92763
ID AAZ92763 standard; DNA; 2076 BP.

XX AC AAZ92763;
XX DT 19-JUN-2000 (first entry)
XX DE Artemisia annua amorphadiene synthase DNA clone.
XX DE Artemisia annua.
XX KW Amorpha-4,11-diene synthase; amorphadiene synthase; FPP;
XX KW farnesyl pyrophosphate; cyclisation; artemisin precursor; antimalarial;
XX KW sesquiterpene; flavouring; fragrance; ss.
XX OS Artemisia annua.
XX PH Key Location/Qualifiers
XX FT CDS 6..1646
XX FT /*tag= a
XX FT /product= "Artemisia annua amorphadiene synthase"

XX EP982404-A1.
XX PD 01-MAR-2000.
XX PF 27-AUG-1998; 98EP-00202854.
XX PR 27-AUG-1998; 98EP-00202854.
XX PA (WALL/) WALLAART T E.
XX PI Wallaart TED, Bouwmeester HJ;
XX DR WPI; 2000-258617/23.
XX DR P-PSDB; AAY81272.
XX PT New isolated DNA sequences and polypeptides comprising amorpha-4,11-diene synthase activity, useful for production of amorphadiene and/or artemisinin.
XX PS Disclosure; Page 10-14; 41pp; English.
XX CC The invention relates to DNA encoding Artemisia annua amorpha-4,11-diene synthase (AAZ92754). This enzyme catalyses the cyclisation of farnesyl pyrophosphate (FPP), a ubiquitous precursor of primary and secondary metabolites, into amorpha-4,11-diene. Amorphadiene is a precursor of artemisin, a sesquiterpene lactone endoperoxide produced by A. annua which is a promising antimalarial drug candidate. Amorphadiene synthase is a branch point enzyme in the artemisin biosynthetic pathway, and the cyclisation of FPP is therefore likely to be the rate-limiting step. Artemisin is present in very low concentrations in A. annua, making it expensive for use as an antimalarial drug, and attempts at organic synthesis have been unsuccessful. Sesquiterpenes such as amorphadiene are also useful as flavour and fragrance compounds in the food and perfume industries. Terpenes also play a role in plant-insect interactions, such as the attraction or repulsion of insects by plants. In addition, dihydroartemisinic acid, an intermediate in the metabolic route from amorphadiene to artemisinin in A. annua, can be used as an antioxidant. DNA encoding amorphadiene synthase may be used to generate transgenic plants able to synthesise amorphadiene synthase. Preferably, the plants used for amorphadiene production are plants which naturally produce sesquiterpenes, as these plants already have the basic sesquiterpene synthetic pathways and storage compartments. A. annua may be transformed with amorphadiene synthase expression constructs, enabling more amorphadiene to be produced and hence increasing artemisin production. It may then be economically feasible to extract artemisin from such plants for use as an antimalarial. The present sequence represents a DNA clone encoding A. annua amorphadiene synthase, which is not referred to further in the specification
XX SQ Sequence 2076 BP; 681 A; 363 C; 412 G; 620 T; 0 U; 0 Other;

Query Match 99.3%; Score 1637.4; DB 3; Length 2076;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1641; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 2 CATGGCACCCTTACAGAGAAAAACCTATTGCCCCCATTCGCCAATTTCTCCCAAGCATTTG 61
Db 5 CATGTCACTTACAGAGAAAAACCTATTGCCCCCATTCGCCAATTTCTCCCAAGCATTTG 64
QY 62 GGGAGATCAGTTTCTCTCATCTATCAAAAGCAAGTAGAGCAAGGGGTGGAAACAGATAGTGA 121
Db 65 GGGAGATCAGTTTCTCTCATCTATCAAAAGCAAGTAGAGCAAGGGGTGGAAACAGATAGTGA 124
QY 122 TGATTTAAAAAGAGAGTGGCGCAACTACTATAAAGAGAGCTTTGGATATTTCTTATCAACA 181
Db 125 TGATTTAAAAAGAGAGTGGCGCAACTACTATAAAGAGAGCTTTGGATATTTCTTATCAACA 184
QY 182 TGCCAAATTTCTTGAAGCTGATTGATGAATTCACACGCTTGGAAATACCGTATCATCTTGA 241
Db 185 TGCCAAATTTCTTGAAGCTGATTGATGAATTCACACGCTTGGAAATACCGTATCATCTTGA 244
QY 242 ACGGGAGATGATCATGCAATTCGCAATGTATTTATTAAGAAACATATGCTGTAATACTGGAATG 301

245	Db		ACGGGAGATTGATCATGCATTCGCAATGATATTTATGAAACATATGCTGATCAACTCGGAATGG	304
302	QY		TGACCGCTCTTCCTTATGTTCCGTCCTTATGCGAAAGCAAGGATATTATGTTACATCTGA	361
305	Db		TGACCGCTCTTCCTTATGTTCCGTCCTTATGCGAAAGCAAGGATATTATGTTACATCTGA	364
362	QY		TGTTTTCAATPAACTATAAAGACAAAAATGGAGCGTTCAAGCAATCGTTAGCTAATGATGT	421
365	Db		TGTTTTCAATPAACTATAAAGACAAAAATGGAGCGTTCAAGCAATCGTTAGCTAATGATGT	424
422	QY		TGAAGTTTGCTTGAGTTGTACGAGCAACTTCTATGAGGGTACCTGGGAGATTATATT	481
425	Db		TGAAGTTTGCTTGAGTTGTACGAGCAACTTCTATGAGGGTACCTGGGAGATTATATT	484
482	QY		AGAAGATGCTCTTTGGTTTACACGATCTCGTCTTAGCATTATGACAAAAAGATGCTTTTC	541
485	Db		AGAAGATGCTCTTTGGTTTACACGATCTCGTCTTAGCATTATGACAAAAAGATGCTTTTC	544
542	QY		TACAAACCCCGCTCTTTTACCGAAATACAAACGGGCATAAGCAACCCCTTTGGAAAG	601
545	Db		TACAAACCCCGCTCTTTTACCGAAATACAAACGGGCATAAGCAACCCCTTTGGAAAG	604
602	QY		GTTTGCCNAGATAGAGCGGCGGACATTCCTTTCTATCAACAAAGATTTCTATAA	661
605	Db		GTTTGCCNAGATAGAGCGGCGGACATTCCTTTCTATCAACAAAGATTTCTATAA	664
662	QY		CAAGACTTTACTTTAAACTTCTTAAGTTAGAGTTCAATTTGCTTCAGTTCATTGCACAAGGA	721
665	Db		CAAGACTTTACTTTAAACTTCTTAAGTTAGAGTTCAATTTGCTTCAGTTCATTGCACAAGGA	724
722	QY		AGAGCTCAGCCATGTGTGCAAAATGGTGGAAAGCTTTCGATATCAAGAAAGAACGCACCTTG	781
725	Db		AGAGCTCAGCCATGTGTGCAAAATGGTGGAAAGCTTTCGATATCAAGAAAGAACGCACCTTG	784
782	QY		TTTAAAGAGATAGAAATTTGTTGAATGTCTACTTTTGGGACCTAGGTTTCAGGCTATGAGCCACA	841
785	Db		TTTAAAGAGATAGGGTTGTTGAATGTCTACTTTTGGGACCTAGGTTTCAGGCTATGAGCCACA	844
842	QY		GTATTCGGGGCTAGAGTTTTCTTCAAAAGCTGTTGCTTTATTACTCTTATAGATGA	901
845	Db		GTATTCGGGGCTAGAGTTTTCTTCAAAAGCTGTTGCTTTATTACTCTTATAGATGA	904
902	QY		CACCTATGATGCGTATGGTACTTTATGAAGAACTTAAAGATCTTTACTCAAGCTGTGAAAG	961
905	Db		CACCTATGATGCGTATGGTACTTTATGAAGAACTTAAAGATCTTTACTCAAGCTGTGAAAG	964
962	QY		GTGGTCAATTACATGCTTAGACACACCTTCAGAAATACATGAAACCGGATATACAAATPATT	1021
965	Db		GTGGTCAATTACATGCTTAGACACACCTTCAGAAATACATGAAACCGGATATACAAATPATT	1024
1022	QY		CATGGATACATACACAGAAATGGAAGAAATTTCTTGAAAGAGGGAAGACAGATCTATT	1081
1025	Db		CATGGATACATACACAGAAATGGAAGAAATTTCTTGAAAGAGGGAAGACAGATCTATT	1084
1082	QY		TAACTCGGGCAAGAAATTTGCTGAAAGAGTTTGTGTAGAAAACCTGATGTTGGAAGCAAAATG	1141
1085	Db		TAACTCGGGCAAGAAATTTGCTGAAAGAGTTTGTGTAGAAAACCTGATGTTGGAAGCAAAATG	1144
1142	QY		GGCAATGAGGGACACATACCAACCACTGAAGAGCATGATCCAGTTGTAATCATTTACTGG	1201
1145	Db		GGCAATGAGGGACACATACCAACCACTGAAGAGCATGATCCAGTTGTAATCATTTACTGG	1204
1202	QY		CGGTGCTAACTGTTACAAACACTTGTATCTTGGCATGAGTGATATATTTCAAAAAGA	1261
1205	Db		CGGTGCTAACTGTTACAAACACTTGTATCTTGGCATGAGTGATATATTTCAAAAAGA	1264
1262	QY		GTCTGTGGAATGGCTGCTCTGCACCTCTCTTTTGTAGATACTTCAGGTATACTTGTGTCG	1321
1265	Db		GTCTGTGGAATGGCTGCTCTGCACCTCTCTTTTGTAGATACTTCAGGTATACTTGTGTCG	1324
1322	QY		ACGCCATAAATGATCTCATGACCCCAAGGCGGACGAAGAAAGAAAAATAGTTTCATCGAG	1381

Db	1325	ACGCCTAAATGATCTCATGACCCCAAGGCCGAGCAAGAAAGAAAACAATAGTTTCATCGAG	1388
Qy	1382	CCTTTGAAGAGTTATATGAAGGAATATATATGTCAATGAGGAGTATGCCCAACCTTGATTTA	1441
Db	1385	CCTTTGAAGAGTTATATGAAGGAATATATGTCAATGAGGAGTATGCCCAACCTTGATTTA	1444
Qy	1442	CAAGGAAAGTAGAAGATGTGTGGAAAGATATAAACCCGAGAGTACCTCACAACCTAAAACAT	1501
Db	1445	CAAGGAAAGTAGAAGATGTGTGGAAAGATATAAACCCGAGAGTACCTCACAACCTAAAACAT	1504
Qy	1502	TCCAAAGCCGCTTATTCATGCGCTGTGATCTATTTGTGCGCAGTTTCTTTGAGGTTCAATATGC	1561
Db	1505	TCCAAAGCCGCTTATTCATGCGCTGTGATCTATTTGTGCGCAGTTTCTTTGAGGTTCAATATGC	1564
Qy	1562	AGGAAAGGATAACTTCCACACGTATGGGAGACCAATACAAACATCTCTATAAAGTCTCTACT	1621
Db	1565	AGGAAAGGATAACTTCCACACGTATGGGAGACCAATACAAACATCTCTATAAAGTCTCTACT	1624
Qy	1622	CGTTTATCTTATGAGTATATGAGGATC	1648
Db	1625	CGTTTATCTTATGAGTATATGACTACC	1651
RESULT 4			
AAZ92757			
XX	ID	AAZ92757 standard; cDNA; 538 BP.	
AC	AAZ92757;		
XX	19-JUN-2000	(first entry)	
DT	Artemisia annua amorphadiene synthase 538 bp cDNA probe.		
DE	Amorpha-4,11-diene synthase; amorphadiene synthase; FPP;		
KW	farnesyl pyrophosphate; cyclisation; artemisin precursor; antimalarial;		
KW	sesquiterpene; flavouring; fragrance; hybridisation probe; ss.		
XX	Artemisia annua.		
OS	EP982404-A1.		
XX	01-MAR-2000.		
PD	27-AUG-1998; 98BP-00202854.		
PF	27-AUG-1998; 98BP-00202854.		
XX	(WALL/) WALLAART T E.		
XX	Wallaart TED, Bouwmeester HJ;		
PI	WPI; 2000-258617/23.		
DR	P-PSDB; AAY81720.		
DR	New isolated DNA sequences and polypeptides comprising amorpha-4,11-diene		
PT	synthase activity, useful for production of amorphadiene and/or		
PT	artemisinin.		
XX	Example 2; Fig 7; 4lpp; English.		
PS	The invention relates to DNA encoding Artemisia annua amorpha-4,11-diene		
XX	synthase (AAZ92754). This enzyme catalyses the cyclisation of farnesyl		
CC	pyrophosphate (FPP), a ubiquitous precursor of primary and secondary		
CC	metabolites, into amorpha-4,11-diene. Amorphadiene is a precursor of		
CC	artemisin, a sesquiterpene lactone endoperoxide produced by A. annua		
CC	which is a promising antimalarial drug candidate. Amorphadiene synthase		
CC	is a branch point enzyme in the artemisin biosynthetic pathway, and the		
CC	cyclisation of FPP is therefore likely to be the rate-limiting step.		
CC	Artemisin is present in very low concentrations in A. annua, making it		
CC	expensive for use as an antimalarial drug, and attempts at organic		
CC	synthesis have been unsuccessful. Sesquiterpenes such as amorphadiene are		
CC	also useful as flavour and fragrance compounds in the food and perfume		
CC	industries. Terpenes also play a role in plant-insect interactions, such		

CC as the attraction or repulsion of insects by plants. In addition,
CC dihydroxycinnamic acid, an intermediate in the metabolic route from
CC amorphadiene to artemisinin in A. annua, can be used as an antioxidant.
CC DNA encoding amorphadiene synthase may be used to generate transgenic
CC plants able to synthesize amorphadiene synthase. Preferably, the plants
CC used for amorphadiene production are plants which naturally produce
CC sesquiterpenes, as these plants already have the basic sesquiterpene
CC synthetic pathways and storage compartments. A. annua may be transformed
CC with amorphadiene synthase expression constructs, enabling more
CC amorphadiene to be produced and hence increasing artemisinin production. It
CC may then be economically feasible to extract artemisinin from such plants
CC for use as an antimalarial. The present sequence represents a 538 bp
CC amorphadiene synthase probe isolated using degenerate PCR primers
CC AA292755- AA292756 from cDNA generated from RNA from A. annua which had
CC previously been subjected to stress conditions. This probe was used to
CC screen an A. annua cDNA library for clones encoding amorphadiene synthase
XX
XX
SQ Sequence 538 BP; 160 A; 102 C; 119 G; 157 T; 0 U; 0 Other;

Query Match 31.0%; Score 510.6; DB 3; Length 538;
Best Local Similarity 97.4%; Pred. No. 4.1e-127;
Matches 519; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 381 GACAAATGGAGCGTTCAAGCAATCGTTAGCTTAATGATGTTCAAGGTTTGGTTGAGTTG 440
Db 1 GATGAGATGGGAATTTAAGGAATCGTTAGCTTAATGATGTTGAAGTTTGGTTGAGTTG 60
QY 441 TACGAAGCAACTTCTATGAGGTACTGGGGATTTATATAGAGATGCTCTTGGTTT 500
Db 61 TACGAAGCAACTTCTATGAGGTACTGGGGATTTATATAGAGATGCTCTTGGTTT 120
QY 501 ACAGATCTCGTCTTAGCATTTATGACAAAGATGCTTTTCTCAAAACCCCGCTCTTTT 560
Db 121 ACACGATCTCGTCTTAGCATTTATGACAAAGATGCTTTTCTCAAAACCCCGCTCTTTT 180
QY 561 ACCGAAATACACGGGCACCTAAGCAACCCCTTTGGAAAGGTTGCCAAGATAGAGGCG 620
Db 181 ACCGAAATACACGGGCACCTAAGCAACCCCTTTGGAAAGGTTGCCAAGATAGAGGCG 240
QY 621 GCGAGTACATCTCTTCTATCAACCAAGATTTCTAACAAGACTTTTACTTAAACTT 680
Db 241 GCGAGTACATCTCTTCTATCAACCAAGATTTCTAACAAGACTTTTACTTAAACTT 300
QY 681 GCTAGTTAGATTCAATTTGCTTTCATGCTATTCGACAAAGAGCTCAGCCATGTGTGC 740
Db 301 GCTAGTTAGATTCAATTTGCTTTCATGCTATTCGACAAAGAGCTCAGCCATGTGTGC 360
QY 741 AATGTTGGAAGCTTTCGATATCAAGAGAACGACCTTGTATTAGAGATAGATTGTT 800
Db 361 AATGTTGGAAGCTTTCGATATCAAGAGAACGACCTTGTATTAGAGATAGATTGTT 420
QY 801 GAATGCTACTTTTGGGACTAGGTTCAAGCTATGAGCCACAGTATTCCTGGGCTAGAGTT 860
Db 421 GAATGCTACTTTTGGGACTAGGTTCAAGCTATGAGCCACAGTATTCCTGGGCTAGAGTT 480
QY 861 TTCTTCACAAAGCTGTTGCTTATTAACCTTTATAGATGACACTTATGATGC 913
Db 481 TTCTTCACAAAGCTGTTGCTTATTAACCTTTATAGATGACACTTATGATGC 533

RESULT 5
ADE82707
ID ADE82707 standard; DNA; 1973 BP.
XX
XX ADE82707;
XX
XX 29-JAN-2004 (first entry)
XX
XX Terpenoid biosynthesis related H64 strawberry DNA #16.
XX
XX isoprenoid; bio-active compound synthesis; pesticide; dermatological;
KW cytostatic; immunosuppressive; virucide; flavour; fragrance;
KW bio-control agent; food additive; food industry; pest control;

KW degreasing solvent; plasticizer; dye carrier; dental caries;
KW dental plaque; skin disorder; immunosuppressive; anti-leukaemia;
KW anti-retroviral; monoterpene alcohol linalool;
KW sesquiterpene alcohol nerolidol; monoterpene; strawberry; gene; ds.
XX
XX Fragaria x ananassa.
XX
XX EP1231273-A1.
XX
XX 14-AUG-2002.
XX
XX 12-FEB-2001; 2001EP-00200488.
XX
XX 12-FEB-2001; 2001EP-00200488.
XX
XX (PLAN-) PLANT RES INT BV.
XX
XX Aharoni A, Verhoeven HA, Jongsma MA, Bouwmeester HJ;
XX
XX WPI; 2003-879727/82.
XX
XX Novel recombinant nucleic acid encoding proteinaceous molecule, useful
XX for producing flavor, fragrance and/or biocontrol agent which is useful
XX as food additive in processed food industry and as antimicrobial agent.
XX
XX Disclosure; Page; 52pp; English.

CC The invention relates to a novel isolated or recombinant nucleic acid or
CC its functional fragment, encoding a proteinaceous molecule essentially
CC capable of isoprenoid bio-active compound synthesis when provided with a
CC suitable substrate under appropriate reaction conditions. The novel
CC recombinant isoprenoid bio-active compound synthesis nucleic acid and its
CC protein have the following activities: pesticide, dermatological,
CC cytostatic, immunosuppressive, and virucide. The novel recombinant
CC isoprenoid bio-active compound synthesis nucleic acid is useful for
CC producing flavour, fragrance, and/or a bio-control agent, by transforming
CC or transfecting a suitable host with the recombinant isoprenoid bio-
CC active compound synthesis nucleic acid, expressing the recombinant
CC isoprenoid bio-active compound synthesis nucleic acid in the presence of
CC a suitable substrate, and optionally isolating the formed product. The
CC bio-control agent is useful as an anti-microbial agent, as a food
CC additive in the processed food industry to modify the taste of syrups,
CC ice-creams, frozen desserts, yogurts, confectionery and like products, as
CC a flavouring agent for oral medications and vitamins, and for providing
CC additional flavour/aroma in beverages, including alcoholic beverages. The
CC bio-control agent is also useful for enhancing or reducing flavour,
CC aroma, fragrance or scent of plants, natural products, and/or synthetic
CC or artificial products, and for the industrial synthesis of nature
CC identical flavour/aroma substances and/or artificial flavour/aroma
CC substances. The bio-control agent is also useful as a pest control agent
CC for the biological control of the interaction between plants and insects
CC and/or plants and microorganisms, for providing flavour/aroma in
CC cosmetics, creams, sun-protectant products, hair conditioners, cleaning
CC products, personal care products and health care products, as a
CC disinfectant additive and in the preparation of a composition. The novel
CC recombinant isoprenoid bio-active compound synthesis nucleic acid or its
CC fragments is useful as a molecular marker or diagnostic tool. The protein
CC of the novel recombinant isoprenoid bio-active compound synthesis nucleic
CC acid is useful for the production of an antagonist e.g. an antibody or
CC its functional equivalent which is useful for inhibiting the synthesis of
CC the bio-control agent. A composition, containing the bio-control agent,
CC is a pharmaceutical or nutraceutical, useful for augmenting or enhancing
CC the aroma and/or taste of food or non-food products, and/or protection of
CC food or non-food products against fungal contamination and/or pest
CC infestation. The composition is also useful for the biological control of
CC pests, for the protection of stored products and for the prevention or
CC treatment of disease. The bio-control agent is useful as a degreasing
CC solvent, plasticizer and dye carrier. The composition is useful for
CC replacing potentially carcinogenic synthetic food additives currently
CC used. The composition is also useful for treating dental caries, dental
CC plaque and skin disorders, and for immunosuppressive, anti-leukaemia and
CC anti-retroviral treatment. The novel recombinant isoprenoid bio-active
CC compound synthesis nucleic acid or its protein is useful for the

CC synthesis of monoterpene alcohol linalool and sesquiterpene alcohol
 CC nerolidol, and monoterpenoid. This polynucleotide sequence represents the
 CC DNA encoding an H64 protein used in the terpenoid biosynthesis method of
 CC the invention. NOTE: This sequence is not shown in the specification. It
 CC has been obtained from electronic data supplied with this specification
 CC from the European Patent Office.

XX
 SQ Sequence 1973 BP; 612 A; 361 C; 434 G; 566 T; 0 U; 0 Other;
 Query Match 24.0%; Score 396.2; DB 9; Length 1973;
 Best Local Similarity 53.8%; Pred. No. 4.4e-96;
 Matches 877; Conservative 0; Mismatches 733; Indels 21; Gaps 2;

QY	22	AACCTATTGCGCCCACTTCCCAAGCAATTTGGGGAGATCAGATTTCTCATCT	81
DB	53	AAGTTGTTGCGGCACACAAATTTAACTAGCTTTGGGGAGATCGTTTGTAACT	112
QY	82	ATCAAAAGCAAGTAGACGAGGGTGGACAGATAGTGAATGATTTAAAAAAGAGTGC	141
DB	113	ATGCCGAAGACATTATAACTCAACTCAATGCAAGAACAACTGAGGAGCTGAAACAAG	172
QY	142	GGCAACTACTAAAGAGCTTTGGATATTCCTATGAACATGCCAATTTGTTGAAGCTGA	201
DB	173	TAGTGAGGAAGAGATTTCACTAATGCTGCTGATGATTTCTCACTCAACTGAAGCTAA	232
QY	202	TTGATGAAATTCACGCCCTTGGAAATACCGTATCACTTTGAAACGGAGATTGATCGCAT	261
DB	233	TTGATGAATCCAGCGCTCGGTGGCTTACCATTTCGAAGCGAATAGATCAAGCCC	292
QY	262	TGCAATGATTTTGAACAATATGATGATGATTAACCTGGAATGGTG-----ACCGCTCTT	312
DB	293	TGGAACGATATACATGAGACATATCAAGATATTCATGATGGTGGTGTATCTGTACAATGTTG	352
QY	313	CTTATGTTTCGGCTTTATGCGAAGCAAGGATATTTATGTTACATGTGATGTTTCAATA	372
DB	353	CTCTTGGTTTCGGCTACTCAGCGGACATGGATATAATGTTTCTCGGATGATTCAACA	412
QY	373	ACTATAAGACAAATAATGGAGGTTTCAAGCAATCGTTAGCTAATGATTTGAAGGTTTGC	432
DB	413	AGTTCAAGATACATATGTTGACTACAGAAAGCTTGTCTACTGATCTTTCTGGTATGC	472
QY	433	TTGAGTTGTAAGAGCAACTTCTATGAGGTTACCTGGGGAGATTTATTTAGAAGATGCTC	492
DB	473	TGAGCTTTTATGAGCGGCCCATCTCGAGGTTGCATGGAGAAAATTTACTTTGAAGAGCTC	532
QY	493	TTGCTTTTACAGATCTGCTTACATATGACAAAGATGCTTTTCTACAAACCCCG	552
DB	533	TGGTTTTCACCACTCATCTCC-----AGTCAGCAAGTGCANAAAGCTCTT	580
QY	553	CTCTTTTACCGAATACAAACGGGCATTAAGCAACCCCTTTTGGAAAAGGTTGCCAAGAA	612
DB	581	TGCTGAACACAAATACTGAAGCGGTAGAGAGCTTCTACTAAAACATTTGGAGAGT	640
QY	613	TAGAGCGCGCGATACATTCCTTTCTATCAACAAAGATTTCTCATACAGACTTTTAC	672
DB	641	TAGGTGCTCGCGGTTACATGTCAATATATCAAGATGAAGCTTTCATACAGTGAATAATAC	700
QY	673	TTAACTTGTAGTTAGATTCATTTGCTTCAGTTCATTTGACAGGAGAGCTCAGCC	732
DB	701	TGAAACTTGCANAAATTTGATTTTAATGTTTTCAGTGTTCACAAAAGGAACTCAGTG	760
QY	733	ATGTGTCAAAATGGTGAAGCTTTTCGATATCAAGAAACGACCTTCTTTTAAGAGATA	792
DB	761	ACATTTCAAGATGGTACAGGAACCTGACTTTTGAAGGAGGATGCTTTTGTCTCGAGATA	820
QY	793	GAATTTGTAATGCTACTTTTGGGACTAGGTTTCAGGCTATGAGCCACAGTATTCGCGGG	852
DB	821	GGATCGTGGAGTTGTTCTTTTGGATAGCAGGAATATTTTCGAACCTGTAATAGCTTTTG	880
QY	853	CTAGAGTTTCTTCAAAAGCTGTTGCTGTTTATTAATCTTTATAGATGACACTTATGATG	912
DB	881	GGAGACACATTTCTGACTAACTGATTGAGATAACACAGTAATGGATGATATGTATGATG	940

QY	913	CGTATGCTACTTATGAAGAACTTAAAGATCTTTACTGAAGCTGTTGAAAGGTGGTCAATTA	972
DB	941	CATTGCGTACATTCGAAGAACTCGTCACTTGAAGCAATTCGACAGGTGGGATGCAA	1000
QY	973	CATGCTTAGACACATTCGAGATACATGAACCGATATACAAATTTATTCATGATACAT	1032
DB	1001	GTTGCATGGATCAACTGCCAGACTATATGCAACCATTTTATATACACACTTCTTGGATGTTA	1060
QY	1033	ACACAGAAATGGAAAGATTTCTTCAAGGAGGGAAGAAACAGATCTATTTAACTCGCGCA	1092
DB	1061	TCGATGAAGTTGAAGGAGCTGCAAGGAGAGATCTTCGGAATTCATCTAGGCA	1120
QY	1093	AAGAAATTTGAAAGAGTTTGTAGAAAACCTGATGGTGAAGCAAAATGGGCAAAATGAGG	1152
DB	1121	AAGAAATTTGAAAGATCAAGCCAGGCTCTACTTCGCTGAGGCCATATGGTTCCACGAAG	1180
QY	1153	GACACATACCAACCTGAAGAGCATGATCCAGTTGTAATCATTTACTGCGGCTGCTAAC	1212
DB	1181	GATGCCACCCCAAAATGGATGGGTATATGCGAGTTGCGGCATCTTCTGCTGGAACACCA	1240
QY	1213	TGCTTACCAACAACTTGTATCTTGGCATGATGATATATTCACAAAAGAGTCTGTGCAAT	1272
DB	1241	TGCTTTCCGCTGCTGCTTTTAGTAGGCTAGGAGATTAACCAAAATTTGAAATTCGAGT	1300
QY	1273	GGCTGTCTCTGCACTCTCTTTTATAGATCTCAGGTATATCTTGGTTCGACGCCCTAAATG	1332
DB	1301	GGCTGACCAATGAGCTTAAATCCCTTAGAGCTTCGAATACCATATTTAGGCTTATGGATG	1360
QY	1333	ATCTATGACCCCAAGCCGAGCAAGAAAGAAACATAGTTTCATCGAGCCCTTGAAGTT	1392
DB	1361	ACATTCCTGGGTACAAATTTGAGAAAGAGAGAGGCGCATGTTGCTTCTAGTATTTGATGCT	1420
QY	1393	ATATGAAGGAATATATGTCAATGAGGAGTATGCCCAACCTTGATTTTACAAGGAAGTAG	1452
DB	1421	ACATGAATGAATAGGGGTTTCAGAGCAAGAGACAAATGTATCTTCAACAAACGAATTG	1480
QY	1453	AAGATGTGTGAAAGATATAAACCGAGAGTACCTCAACCTPAAACAAATTCAGAGCCGT	1512
DB	1481	TGGATTCGTGGAAGGATATAACGAAGAGTTTCTGAGACCCACTGCTGCTCCAGTCCCTG	1540
QY	1513	TATGTAGGCTGTGATCTATTTGTCAGCTTCTTGAAGTTCAATATTCAGAAAGGATA	1572
DB	1541	TGCTTAATCGTGTCTTAACTAACCCGAGTGGTGTGCTGCTTTACAAAAGGGGAGATG	1600
QY	1573	ACTTCACACTATGGGAGAGCAATACAAACATCTCATAAAGTCTCTACTCGTTTATCTTA	1632
DB	1601	CTTTCAGCAATGTCGGAATACTGATGAAGATTTGTTGTCGAATGTTTATGATCCAG	1660
QY	1633	TGAGTATATGA 1643	
DB	1661	TGCCACTCTGA 1671	

RESULT 6
 ADE82704
 ID ADE82704 standard; DNA; 2605 BP.
 XX
 AC ADE82704;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Terpenoid biosynthesis related H64 strawberry DNA #15.
 XX
 KW isoprenoid; bio-active compound synthesis; pesticide; dermatological;
 KW cytoskeletal; immunosuppressive; virucide; flavour; fragrance;
 KW bio-control agent; food additive; food industry; pest control;
 KW degreasing solvent; plasticizer; dye carrier; dental caries;
 KW dental plaque; skin disorder; immunosuppressive; anti-leukaemia;
 KW anti-retroviral; monoterpene alcohol linalool;
 KW sesquiterpene alcohol nerolidol; monoterpene; strawberry; gene; de.
 XX
 OS Fragaria x ananassa.
 XX

EP1231273-A1.
 14-AUG-2002.
 12-FEB-2001; 2001EP-00200488.
 12-FEB-2001; 2001EP-00200488.
 (PLAN-) PLANT RES INT BV,
 Aharoni A, Verhoeven HA, Jongsma MA, Bouwmesster HU;
 WFI; 2003-879727/82.
 Novel recombinant nucleic acid encoding proteinaceous molecule, useful for producing flavor, fragrance and/or biocontrol agent which is useful as food additive in processed food industry and as antimicrobial agent.
 Disclosure; Page: 52pp; English.
 The invention relates to a novel isolated or recombinant nucleic acid or its functional fragment, encoding a proteinaceous molecule essentially capable of isoprenoid bio-active compound synthesis when provided with a suitable substrate under appropriate reaction conditions. The novel recombinant isoprenoid bio-active compound synthesis nucleic acid and its protein have the following activities: pesticide, dermatological, cytostatic, immunosuppressive, and virucide. The novel recombinant isoprenoid bio-active compound synthesis nucleic acid is useful for producing flavour, fragrance, and/or a bio-control agent, by transforming or transfecting a suitable host with the recombinant isoprenoid bio-active compound synthesis nucleic acid, expressing the recombinant isoprenoid bio-active compound synthesis nucleic acid in the presence of a suitable substrate, and optionally isolating the formed product. The bio-control agent is useful as an anti-microbial agent, as a food additive in the processed food industry to modify the taste of syrups, ice-creams, frozen desserts, yogurts, confectionery and like products, as a flavouring agent for oral medications and vitamins, and for providing additional flavour/aroma in beverages, including alcoholic beverages. The bio-control agent is also useful for enhancing or reducing flavour, aroma, fragrance or scent of plants, natural products, and/or synthetic or artificial products, and for the industrial synthesis of nature identical flavour/aroma substances and/or artificial flavour/aroma substances. The bio-control agent is also useful as a pest control agent for the biological control of the interaction between plants and insects and/or plants and microorganisms, for providing flavour/aroma in cosmetics, creams, sun-protectant products, hair conditioners, cleaning products, personal care products and health care products, as a disinfectant additive and in the preparation of a composition. The novel recombinant isoprenoid bio-active compound synthesis nucleic acid or its fragments is useful as a molecular marker or diagnostic tool. The protein of the novel recombinant isoprenoid bio-active compound synthesis nucleic acid is useful for the production of an antagonist e.g. an antibody or its functional equivalent which is useful for inhibiting the synthesis of the bio-control agent. A composition, containing the bio-control agent, is a pharmaceutical or nutraceutical, useful for augmenting or enhancing the aroma and/or taste of food or non-food products, and/or protection of food or non-food products against fungal contamination and/or pest infestation. The composition is also useful for the biological control of pests, for the protection of stored products and for the prevention or treatment of disease. The bio-control agent is useful as a degreasing solvent, plasticizer and dye carrier. The composition is useful for replacing potentially carcinogenic synthetic food additives currently used. The composition is also useful for treating dental caries, dental plaque and skin disorders, and for immunosuppressive, anti-leukaemia and anti-retroviral treatment. The novel recombinant isoprenoid bio-active compound synthesis nucleic acid or its protein is useful for the synthesis of monoterpene alcohol linalool and sesquiterpene alcohol nerolidol, and monoterpenoid. This polynucleotide sequence represents the DNA encoding an H64 protein used in the terpenoid biosynthesis method of the invention. NOTE: This sequence is not shown in the specification. It has been obtained from electronic data supplied with this specification from the European Patent Office.

Query Match	Best Local Similarity	Mismatches	Indels	Gaps
22	AACTATATTCGCCCAATGGCAACTTTCTCTCCAAAGCAATTTGGGGAGATCAATTTCTCATCT	81		
744	AACTTGTCTGGCGCAGACAGAAATTTTAAACCTAGCCGTTTGGGAGATCGGTTTGTCTAACT	803		
82	ATCAAAAGCACTAGAGCAAGGGTGGACAGATAGTGAATGATTTTAAAAAAGAGAGTGC	141		
804	ATCCGAAGACATTAATACTCAAACTCAATGCAAGAACAAAGTTGAGGAGCTGAACAAG	863		
142	GGCAACTACTAAAGAAGCTTTGGATATTCTTATGAAACATGCAATTTGTTCGAAGCTGA	201		
864	TAGTGAGGAGGAGATTTCACTAATGCTGCTGATGATTTCTTACATCAACTGAAGCTAA	923		
202	TTGATGAATTCACGCCCTTGGAAATACCGTATCACCTTTGAAACGGGAGATTTGATCGCAT	261		
924	TTGATGAATTCACGCCCTCGGTGTGGCTTACCATTTCGAAAGCGAAATAGATCAAGCCC	983		
262	TGCAATGTATTTATGAACATATGGTGATACTGGAAATGGTG-----ACCGCTCTT	312		
984	TGGAACGTATACATGAGACATATCAAGATATTCAATGATGGTGGTGATCTGTACAATGTTG	1043		
313	CCTTATGGTTCGGCTCTTATGCGAAAGCAAGGATATTATGTTACATGTGATGTTTCAATA	372		
1044	CTCTTGGTTCCTGGCTACTCAGCGACATGGATATAATGTTTCTCGGATGATTTCAACA	1103		
373	ACTATAAGACAAAATGAGCGGTTCAAGCAATCGTTAGCTTAATGATGTTGAAGTTTGC	432		
1104	AGTTCAAGATACATAATGGTGACTACAAGAAAAGCTTGGTCACTGATCTTCTCTGGTATGC	1163		
433	TTGAGTTGTACGACGAACTTCTATGAGGGTACCTGGGAGATTTATTTAGAGATGCTC	492		
1164	TGAGCTTTTATGAGCGGGCCACTCTGAGGGTGATGGGAAAAATTTACTTTGAAGAGGCTC	1223		
493	TTGGTTTACACGATCTCTGCTTAGCATTTATGACAAAAGATGCTTTTCTTACAAAACCCCG	552		
1224	TGGTTTTTACCACTCATCTCC-----AGTCGCAAGTGCAAAAGCTCTT	1271		
553	CTCTTTTTCACGAAATACAAACGGGCACTAAAGCAACCCCTTTGGAAAAAGTTGCCAAGAA	612		
1272	TGCTGAAAAACAAATAAATCTGAAGCCGTAGAGAGACTACT---AAAAACTATGGAGAGGT	1328		
613	TAGAGCGGGCGAGTACATTCCTTCTATCAACAAGATTTCTCATACAGAGCTTTAC	672		
1329	TAGGTCTCGGGCTTACATGTCAATATATCAAGATGAAGCTTCATACAGTGAATAATTTAC	1388		
673	TTTAACTCTGTAAGTTAGAGTTCAATTTGCTTCAGTCAATTGACAAAGGAGAGCTCA--G	730		
1389	TGAACTTGCAAAATTAGAATTTTAATGTTGTCAGTGTTTACAAAAAGGAAGTCAAGTG	1448		
731	CCAATGTGTCAAAATGGTGAAAGCTTTTCGATATCAAGAAAGACGCCCTTTGTTTAAAGAGA	790		
1449	ACATTCCTTAAAGATGATCAAGGAACCTGGACCTTGCAAGGAGATGCCCTTTTGTCTGAGA	1508		
791	TAGAAATGTTGAATGCTACTTTTGGGGACTAGGTTCAAGGCTATGAGCCACAGTATTTCCCG	850		
1509	TAGGATCGTGGAGTTGTTCTTTTGGATAGCAGGAATATATTTCCGAACCTGAAATACGTCCT	1568		
851	GGCTAGAGTTTCTTCAAAAAGCTGCTGCTGTTAACTCTTATAGATGACACTTATGA	910		
1569	TAGGAGACACATTTCTGACTAACTGATGAGATTAACAACAGTAAATGATGATGATGA	1628		
911	TGGGTATGTTACTTATGAAGAACTTAAAGATCTTTTACTGAAGCTGTTTGAAGGTGGTCAAT	970		
1629	TGCATTTCCGTCATTCGAGAACTCGTCAACTTGACTGAAGCAATTTGACAGGTGGATGC	1688		
971	TACATGCTTAGACACACTTCCGAATATACATGAACCGATATACAAATTTATCATGATAC	1030		
1689	AACTTGCATGGATCAACTGCGACACTATATGAACCACTTTTATTTATTTATTTATTTATTT	1748		

```
QY 1031 ATACACGAATGAAGAATTTCTTCAAGGAGGAGAAAGACAGATCTATTTAACTGGCG 1090
Db 1749 TATCGATGAAGTTGAAGAGGAGCTGACAAAGCAAGAGATCTTACCGAATTCACATCGC 1808
QY 1091 CAAGAATTTGTGAAGAGTTGTTAGAAA CTTGATGTTTGAAGCAAAATGGCGAAATGA 1150
Db 1809 AAAAGAAATATGAAGATCAAGCAGGCTCTACTTCGCTGAGGCGAGATGTTCCACGA 1868
QY 1151 GGGACACATACCAACCTGAGAGCATGATCCAGTTGTAATCAATCTACTGGCGTCTAA 1210
Db 1869 AGGATGACCCCAAAATAGTATGATATGCGAGTTGCGGCACTCTTGTGCGTAACAC 1928
QY 1211 CCTGTTACAAACAACCTGTTATCTTTGGCATGATATATTCACAAAAGAGTCTGTCA 1270
Db 1929 CATGCTTCCGTGCTGTTAGTAGGATGGAGACATTAACAAATTTGAATTCGA 1988
QY 1271 ATGGGCTGCTCTGCACTCTCTCTTTTAGATACTCAGGTATCTTGGTGGAGCGCTAAA 1330
Db 1989 GTGGCTGACCAATGAGCCTAAATCTTAGAGCTTCGAATACCATATTTAGGCTTATGA 2048
QY 1331 TGATCTCATGCCACACGCGGCGACAGAAAGAAACATAGTTCTATCGAGCTTGAAG 1390
Db 2049 TGACATTCCTGGGTACAAGTTTGAAGAAAGAGAGGCGCATGTTGCTTCAAGTATTGATG 2108
QY 1391 TTATATGAAGGAATATAATGTCAATGAGGAGTATGCCAAACCTTGATTTACAAGGAAT 1450
Db 2109 CTACATGAATGAATACGGGTTTCAGAGCAAGAGCAATGATATCTTCAACAAACGAT 2168
QY 1451 AGAAGATGTGGAAAGATATAAACCAGAGATACCTCAACAATAAACAATTCACAGGCC 1510
Db 2169 TGTGATTCGTGGAAGGATATAAACAAGAGTTCTGAGACCCACTGCTCCAGTCC 2228
QY 1511 GTTATGATGGCTGATCTATTTTGTGCCAGTTCTTGAAGTTCAATATGCGAGAAAGA 1570
Db 2229 TGTGCTTAATCGTGTCTTAACCTAACCCGAGTGTGATCTGCTTTACAAAGGGGAGA 2288
QY 1571 TAACTTCACAGTATGGGAGAGCAATACAAACATCTCAATAAGTCTCTACTCGTTATCC 1630
Db 2289 TGCCTTCAGCATGTCGGAACACTGATGAAGATTGTATGTGCTCAATGTTATTGATCC 2348
QY 1631 TATGATATATGA 1643
Db 2349 AGTGCCACTCTGA 2361
```

RESULT 7

AA89865 ID AA89865 standard; cdna; 1879 BP.

AC AA89865;

DT 08-OCT-1999 (first entry)

DE L. esculentum germacrene C synthase cdna (insert pLE 20.3).

DE Germacrene C synthase; tomato; transgenic; pathogen resistance;
flavour alteration; neutral value; ss.

OS Lycopersicon esculentum.

PH Key Location/Qualifiers
CDS 39..1685
/*tag= a
/product= "germacrene C synthase"

PN WO9938957-A1.

XX 05-AUG-1999.

XX 02-FEB-1999; 99WO-US002133.

XX 02-FEB-1998; 98US-0073579P.

XX PA (UNIW) UNIV WASHINGTON STATE RES FOUND.
XX PA (REGC) UNIV CALIFORNIA.
XX PI Colby SM, Crook JB, Lemaux PG, Croteau RB;
XX WI; 1999-479180/40.
XX DR P-PSDB; AAY27062.
XX
XX PT New Germacrene C synthase protein and nucleic acid useful for producing
XX PT transgenic plants with increased resistance to pathogens.
XX
XX Claim 12; Fig 6; 60pp; English.
XX
XX This represents the nucleotide sequence of the cdna insert pLE 20.3
XX encoding a Lycopersicon esculentum germacrene C synthase protein. The
XX germacrene C synthase gene is useful for generating transgenic plants
XX with an increased host resistance to pathogens including bacteria, rusts,
XX herbivores, microbes and fungi. Recombinant nucleic acid comprising a
XX specific expression operably to the germacrene synthase gene allows tissue
XX promoter expression of the protein to deter e.g. leaf-eating herbivores,
XX and allows production of plants with an altered flavour and odour
XX profiles to deter e.g. fruit-eating herbivores, or to enhance the plants
XX attractiveness for pollinators, or increase the neutral value of
XX the plant. Polypeptide fragments of the protein are useful as immunogens
XX to raise antibodies, and the polynucleotides are useful as probes or
XX primers for detecting, cloning and confirming alleles and homologues of
XX the germacrene synthase gene
XX
SQ Sequence 1879 BP; 665 A; 312 C; 362 G; 540 T; 0 U; 0 Other;

Query Match 22.1%; Score 364; DB 2; Length 1879;

Best Local Similarity 52.9%; Pred. No. 2e-87;

Matches 865; Conservative 0; Mismatches 745; Indels 24; Gaps 3;

QY 29 TCGCCCATTCGCACATTTCTCCAGCATTTGGGAGATCAGTTTCTCATCTCAAAA 88

Db 65 TCGCCCTTGGCTAAATTTTCCACCATCTGTTGGGATATCATTTCTTATATCTCA 124

QY 89 GCAGTAGACGAGGGTGGAGACAGATAGTGAATGATTTAAAAAGAGTGGCGCACT 148

Db 125 TGAATTAATCAAGAAAGTTGAAGTTGATGATGATCAAGAGACATTTAGAAAAAT 184

QY 149 ACTAAAAAGAGCTTTGGATATTCCTATGAAACATGCCAATTTGTTGAAGCTGATTGATGA 208

Db 185 GCTGGTGAACCTTCGCACATAGCAGCTCAAAA-----GCTTGTGTTGATAGACGC 235

QY 209 AATTCAACGCTTGGATATACCGTATACCTTTGAAGGAGATGATCATGCTTGAATG 268

Db 236 GATGCAACGATTTGGGAGTGGCTTATCATTTCCGATAATGAAATTTGAAACATCCATTCAAAA 295

QY 269 TATTTATGAACAT-----ATGGTGAATACTGGAATGCTGACCGCTCTTCCCT 316

Db 296 CATTTTGTGATCATCGTCCAAACAGAAATGATGATGACAAACCTTTACGTTGTGCTCT 355

QY 317 ATGCTTCCGCTTATGCGAAAGCAAGGATATTTATGTTACATGCTGATGTTTCAATAACTA 376

Db 356 TCGTTTTCGACTGTGAGGCAACAGGCCATTTACATGCTTTCAGATGTTTCAAGCAATT 415

QY 377 TAAAGACAAAATGAGCGTTTCAAGCAATCGTTAGCTAATGATGTTGAAGTTTGTGTA 436

Db 416 CACCAACCAAGATGGGAAATTTCAAGGAAACACTTACTAATGATGTTCCAAAGGATTTATGAG 475

QY 437 GTTGTAGAGCAACTTCTATGAGGTACCTGGGAGATTTATTTAGAGAGTCTCTTGG 496

Db 476 TTTGTATGAAGCATCATCTGAGAGTGGTAATGAGGAGATTTCTTGAAGAGCTCTTAC 535

QY 497 TTTTACACGATCTCTGCTTAGCATTTATGACAAAAGATGCTTTTCTTACAAAACCCGCTCT 556

Db 536 ATTTACCAACCATCTCTCGAGTCTTATGCTCCACTTGGCAATTAATAACTCTCT 595

QY 557 TTTTACCGAAATACAAACGGGCACTTAAAGCAACCCCTTTGGAAAAGGTTCCCAAGATAGA 616

Db 596 TAAGTTGAAGTTGGTGAAGCCCTTAAGTCAAGCTTATTCGATGACTTTACCAAGGATGGG 655
QY 617 GCGGCGCAGTACATTCCTTTCTATCAACAAAGATTTCTATCAAGACTTTTACTTAA 676
Db 656 AGCTAGAAAATACATATCCATTTACGAAAACAATGATGCACACCACTTTGCTTTTGA 715
QY 677 ACTTGTCTAGTGTAGATTCATTTGCTTCTAGTCAATTCGACAAAGAGCTCAGCCATGT 736
Db 716 ATTGTCTAAATGGATTTTAAATGCTGCAAAAGTTTCACAAAGAGCTTAGTATCT 775
QY 737 GTGCAATGCTGAAAGCTTTTCGATATCAAGAAAGCGCACTTGTTTTAAAGAGATAGAAT 796
Db 776 TACAAGGCTGCTGAAAGATTTTGATTTTGCATAATATCCATATCAAGAGACAGTT 835
QY 797 TGTGTAATGCTACTTTTGGGAGTAGGTTGAGGCTATGAGCAAGATTTCCCGGGTAG 856
Db 836 GGTGTAGTGTACTTCTGGAATATTAGAGTGTATTTTGGAGCCAAATATATGCTGCGAG 895
QY 857 AGTTTCTTCAAAAAGCTGCTGCTGTTTAACTTTATAGTACACACTTATGATGGTA 916
Db 896 AAAATGATGACAAAAGTACTCAACCTGACCTCAATTTATGACACACTTTTGATGCTTA 955
QY 917 TGTACTTATGAAGAACTTTAAGATCTTTTACTGAAGCTGTTGAAAGGTGGTCAATTACATG 976
Db 956 TGCAACCTTTGACGAAGCTTGTGAGCTTTCAATGATGCAATCCAGAGATGGGATGCTATGC 1015
QY 977 CTTAGACACTTCCAGATACATGAAACCGATATACAAATTTATTCATGATACATACAC 1036
Db 1016 AATTGATTAATACAAACCATATATGAGACCTGCTTATCAAGCTCTTCTAGACATTTACAG 1075
QY 1037 AGAAATGGAAGATTTCTTCAAGAGGAGGAGCAAGATCTATTTAACTGCGGCAAGA 1096
Db 1076 TGAATGGAACAAGTGTGTTCAAGAGAGTAACTGGACCGGTATATCTATGCAAAAA 1135
QY 1097 ATTGTGAAAGAGTTGTTAGAAACCTGTATGGTTGAAGCAAAATGGGCAAAATGA --GGG 1153
Db 1136 TGAGATGAAAGAGTTGTTGAGAGCTTATTTAAGGAAACCAATGGTTGATGATTTGA 1195
QY 1154 ACATATACCAACCACTGAGAGATGATCAAGTTGTAATCATTTACTGCGGGTGTACCT 1213
Db 1196 CCATATTCAAAATATGAGGAACAAGTGGAGATGCAATGCTGTAAGTCTGCTATATGAT 1255
QY 1214 GCTTACAAACCTGTTATCTTGCATGAGTATATTTACAAAAGAGCTGTGCAATG 1273
Db 1256 GATATCAACACTTGTGCTGCTATGAGAGATTTATATCCACAGACTTTTGAATG 1315
QY 1274 GGTGTCTCTGCACCTCCTCTTTTGTAGATCTCAGGTATATCTTGGTGGAGCCCTTAAATGA 1333
Db 1316 GTTGATGATGATGCTGTGATGCTTGGCTTCCGATTCGTCAGAGCAATGAACGA 1375
QY 1334 TCTATGACCCCAAGCCGAGCAAGAAAGAAATATGTTCTATGAGCCTTGAAAGTTA 1393
Db 1376 TATGTTGGACATGAAGATGAACAAGAAAGAGACATGTAGCTTCACTTATTGAATGTTA 1435
QY 1394 TATGAGGAGATTAATGTCAATGAGAGATGATGCCAAACCTTGATTTTACAGGAAGTGA 1453
Db 1436 CATGAAAGATTTGAGCTTCAAGCAAGAGACTTACATTAAGTTCTGAAAGAGTTCAC 1495
QY 1454 AGATGTGTGGAAGATATAACCGAGAGTACCTCAACAACCTTAAACAAATTCAGAGCCGCT 1513
Db 1496 CAATGATGGAAGGACATAAACAACAACTTCTCCGTCCACTGAAGTACCAATGTTGT 1555
QY 1514 ATTGATGGCTGTGATCTATTTGTCAGCTTTCTTGAAGTTCAATATGCAAGAAAGATAA 1573
Db 1556 CTTGAAACGAGTTCTAAATTTGACACGCTGTGGCTGACACGCTTATATAAGAGAAAGATAC 1615
QY 1574 CTTACACGATGAGGAGACCAATACAAACATCTCATAAAGTCTCTACTCTGTTTCTCTAT 1633
Db 1616 ATATCAAGCCCAAGGAAACCTTAAACATGATTAATTCATATTAATGATCTGT 1675
QY 1634 GAGTATATGAGGAT 1647
Db 1676 CAAAATATAAATAT 1689

RESULT 8

AD82701
ID ADB82701 standard; DNA; 1672 BP.
XX
AC ADB82701;
XX
DT 29-JAN-2004 (first entry)
XX
Terpenoid biosynthesis related H64 strawberry DNA #14.
XX
isoprenoid; bio-active compound synthesis; pesticide; dermatological;
cytostatic; immunosuppressive; virucide; flavour; fragrance;
bio-control agent; food additive; food industry; pest control;
degreasing solvent; plasticizer; dye carrier; dental caries;
dental plaque; skin disorder; immunosuppressive; anti-leukaemia;
anti-retroviral; monoterpene alcohol linalool;
sesquiterpene alcohol nerolidol; monoterpene; strawberry; gene; ds.
XX
OS Fragaria x ananassa.
XX
EP1231273-A1.
XX
14-AUG-2002.
XX
12-FEB-2001; 2001EP-00200488.
XX
12-FEB-2001; 2001EP-00200488.
XX
(PLAN-) PLANT RES INT BV.
XX
Aharoni A, Verhoeven HA, Jongsma MA, Bouwmeester HJ;
XX
WPI; 2003-879727/82.
XX
Novel recombinant nucleic acid encoding proteinaceous molecule, useful
for producing flavor, fragrance and/or biocontrol agent which is useful
as food additive in processed food industry and as antimicrobial agent.
XX
Disclosure; Page; 52pp; English.
XX
The invention relates to a novel isolated or recombinant nucleic acid or
its functional fragment, encoding a proteinaceous molecule essentially
capable of isoprenoid bio-active compound synthesis when provided with a
suitable substrate under appropriate reaction conditions. The novel
recombinant isoprenoid bio-active compound synthesis nucleic acid and its
protein have the following activities: pesticide, dermatological,
cytostatic, immunosuppressive, and virucide. The novel recombinant
isoprenoid bio-active compound synthesis nucleic acid is useful for
producing flavour, fragrance, and/or a bio-control agent, by transforming
or transfecting a suitable host with the recombinant isoprenoid bio-
active compound synthesis nucleic acid, expressing the recombinant
isoprenoid bio-active compound synthesis nucleic acid in the presence of
a suitable substrate, and optionally isolating the formed product. The
bio-control agent is useful as an anti-microbial agent, as a food
additive in the processed food industry to modify the taste of syrups,
ice-creams, frozen desserts, yogurts, confectionery and like products, as
a flavouring agent for oral medications and vitamins, and for providing
additional flavour/aroma in beverages, including alcoholic beverages. The
bio-control agent is also useful for enhancing or reducing flavour,
aroma, fragrance or scent of plants, natural products, and/or synthetic
or artificial products, and for the industrial synthesis of nature
identical flavour/aroma substances and/or artificial flavour/aroma
substances. The bio-control agent is also useful as a pest control agent
for the biological control of the interaction between plants and insects
and/or plants and microorganisms, for providing flavour/aroma in
cosmetics, creams, sun-protectant products, hair conditioners, cleaning
products, personal care products and health care products, as a
disinfectant additive and in the preparation of a composition. The novel
recombinant isoprenoid bio-active compound synthesis nucleic acid or its
fragments is useful as a molecular marker or diagnostic tool. The protein
of the novel recombinant isoprenoid bio-active compound synthesis nucleic

CC acid is useful for the production of an antagonist e.g. an antibody or
CC its functional equivalent which is useful for inhibiting the synthesis of
CC the bio-control agent. A composition, containing the bio-control agent,
CC is a pharmaceutical or nutraceutical, useful for augmenting or enhancing
CC the aroma and/or taste of food or non-food products, and/or protection of
CC food or non-food products against fungal contamination and/or pest
CC infestation. The composition is also useful for the biological control of
CC pests, for the protection of stored products and for the prevention or
CC treatment of disease. The bio-control agent is useful as a degrading
CC solvent, plasticizer and dye carrier. The composition is useful for
CC replacing potentially carcinogenic synthetic food additives currently
CC used. The composition is also useful for treating dental caries, dental
CC plaque and skin disorders, and for immunosuppressive, anti-leukemia and
CC anti-retroviral treatment. The novel recombinant isoprenoid bio-active
CC compound synthesis nucleic acid or its protein is useful for the
CC synthesis of monoterpenic alcohol linalool and sesquiterpene alcohol
CC nerolidol, and monoterpenoid. This polynucleotide sequence represents the
CC DNA encoding an H64 protein used in the terpenoid biosynthesis method of
CC the invention. NOTE: This sequence is not shown in the specification. It
CC has been obtained from electronic data supplied with this specification
CC from the European Patent Office.

XX SQ Sequence 1672 BP; 514 A; 315 C; 385 G; 458 T; 0 U; 0 Other;

Query Match 22.0%; Score 363.6; DB 9; Length 1672;
Best Local Similarity 54.2%; Pred. No. 2.5e-87;
Matches 887; Conservative 0; Mismatches 714; Indels 35; Gaps 6;

QY 22 AACCTATTGCGCCCATTCGCCAATTTCTCCCAAGCATTTGGGGAGATCAGTTTCTCATCT 81
DB 53 AAGTTGTTGGCGGCACAGCAAAATTTAAACCTAGCGTTTGGGGAGATCGGTTTGTCTA 112
QY 82 ATCAAAAGCAAGTAGAGCAAGGGTGGAAACAGATAGTGAATTTAAAAAGAAAGTGC 141
DB 113 AT-----GCCGAGAGACATTAATACTCAAACTCAAAATCAAGAACAAAGTTGAGGAGCTGA 166
QY 142 GGCAACTACTAAAGAGCTTTGGATATCTCTATGAACA---TGCCAAATTTGTTGAAGC 198
DB 167 AACAGTGAAGAGGAAGTATTCATATGCTGCTGATGATTCCTCACTCAACTCAAGTGAAGC 226
QY 199 TGATTTGATGAATTCAGCGCTTGGAATACCGTATACCTTTGAAACGGAGATTTGATCATG 258
DB 227 CAATTTGATGAATCCAGCGCTCGGTGGCTTACCAATTTGAAAGCGAATAGATCAAG 286
QY 259 CATTGCAATGATTTATGAACAATATGTTGATGAATCACTGGAATGGTG-----ACCGCT 309
DB 287 CCTTGGAAACGTATACATGAGACATATCAAGATATTCATGATGGTGGTATCTGTACAATG 346
QY 310 CTTCCTTATGTTCCGCTTATGCGAAGCAAGGATATATGTTACATGTTGATGTTTCA 369
DB 347 TTGCTCTGTTTTCGGCTACTCAGGCGACATGGATATAATGTTTCTCGCGATGATTCA 406
QY 370 ATAACTATAAGACAAATAATGGAGCGTTCAAGCAATCGTTAGCTAATGATGTTGAAGTT 429
DB 407 ACAAGTTCAAGATACATAATGTTGATACAGAAAGCTTGGTCACTGATCTTCTGTTA 466
QY 430 TGCTTGATTTGATGAAGCAACTTCTATGAGGTTACCTGGGAGATTTATTTAGAAGATG 489
DB 467 TGCTGAGCTTTATGAGCGCGCCCATCTGAGGTTGATGGAAGAAAATTTACTTGAAGAGG 526
QY 490 CTCTTGCTTTTACAGATCTGCTTACCATATGACAAAGATGCTTTTCTCAAAACC 549
DB 527 CTCTGTTTTCACCACTCATCTCC-----AGTCAGCAAGTGCAGAAAGCT 574
QY 550 CCGCTCTTTTACCGAAATACAAAGCGGCACTAAAGCAACCCCTTTGGAAAGGTTGCCAA 609
DB 575 CTCTGCTGAAACACAAATACTAGAGCGGTAGAGAGACTACT---AAAAGTATGGAGA 631
QY 610 GAATAGAGCGGCGCAGTACATCTCTTCTATCAACAAGAGTCTCATACAGAGATT 669
DB 632 GGTAGGTCGTCGGGTTTACATGTCAATATATCAAGATGAAGCTTCATACAGTGAATTT 691
QY 670 TACTTAACTTGCTAAGTTAGATTCATTTGCTTCAGTCACTTGCACAAAGGAGCTCA 729

DB 692 TACTGAAACCTGCAAAATTAAGATTTTAAATGTTGTTCAAGTGTATACCAAAAAGGAACCTCA 751
QY 730 --GCCATGTGTGCAAAATGGTGGAAAGCTTTGCATATCAAGAAAGAACGCACTTGTGTTAAG 787
DB 752 GTGACATTCCTTAAGATGGTACAAAGGAATGGACTTTTCAAGGAGGATGCTTTTGTCTCG 811
QY 788 AGATAGAAATTTGTTGAATGCTACTTTTGGGAGCTAGGTTCCAGGCTATGAGCCACAGTATTC 847
DB 812 AGATAGGATCGTGGAGTTGTTCTTTTGGATAGCAGGAATATATTTCGAACCTGGAATACCT 871
QY 848 CCGGGCTAGAGTTTCTTTCACAAAAGCTGTTGTTGTTTATACTCTTATAGATGACACTT 907
DB 872 CTTTAGAGACACATCTGACTAACTGATGATGATGATGATGATGATGATGATGATGATGAT 931
QY 908 TGATGCTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 967
DB 932 TGATGCTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 991
QY 968 AATTACATGTTTACACACTTCCAGAAATACATGAAACCGATATACAAATTTATTCATGGA 1027
DB 992 TGCAGTTTGCATGGATCAACTGCCAGACTATATGCAACCACTTTTATATATATACACTTCT 1051
QY 1028 TACATACACAGAAATGGAAGAATTTCTTGCAAGGAGGAGAAAGACAGATCTATTTAACTG 1087
DB 1052 TGTATGCTGATGAGTTGAGAGGAGCTGCAAGCAAGCAAGAAAGATCTTACCGAATTCATTA 1111
QY 1088 CGCAAAAGAAATTTGTGAAAGAGTTTGTGAAACCTGATGTTGAAAGCAAAATGGGCAAA 1147
DB 1112 CGCAAAAGAAATTTATGAAGAATCAAGCCAGCTCTACTTCCGCTGAGGCCAGATGTTCCA 1171
QY 1148 TGAGGAGACATACCAACCACTGAAAGAGATGATCCAGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1207
DB 1172 CGAAGGATGACCCCAAAATGAGATGATATGCGAGTTTCCGCGATCTTCTGTCGGTAA 1231
QY 1208 TACCTGCTTTACAAACCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1267
DB 1232 CACCATGCTTTCGTCGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1291
QY 1268 CGAATGGCTGCTCTGCTGCTGCT 1327
DB 1292 CGAGTGGCTGACCAATGAGCTTAAATCCCTTAGAGCTTCGAATACCATATTTAGGCTTAT 1351
QY 1328 AATGATCTCATGCCACCAAGCCGAGCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1387
DB 1352 GGATGACATTTGCTGGGTACAAAGTTTGAGAAAGAGAGAGAGGAGGATGTTGCTTCAAGT 1411
QY 1388 AAGTTATATGAAGAAATATAATGTTCAATGAGGAGTATGCCAAACCTTGTATTTTCAAGGA 1447
DB 1412 TTGCTACATGAATGAATACCGGGTTTTCAGAGCAAGAGACAAATGATATCTTCAACAAAG 1471
QY 1448 AGTAGAAGATGTTGGAAGATATAAACCCGAGAGTACCTCACAACTAAAAACATTTCCAAAG 1507
DB 1472 AATTGTTGATTTGTTGGAAGGATATAAACGAAAGAGTTTCTGAGACCCACTGCTGCTCCAGT 1531
QY 1508 GCGGTTATGATGGCTGTTGATCTATTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1567
DB 1532 CCCTGCTGCTTAATGCTGCTTCTTAACCTTAACCCGAGTGGTTGATCTGCTTTTCAAAAGGG 1591
QY 1568 GGATAACTTACACAGTATGGGAGACGAATACAAACATCTCATAAAGTCTCTACTGTTTATA 1627
DB 1592 AGATGCTTTCACCGATGTCGAAACATGATGAAAGATTTGTTGCTGCAATGTTTATTTGA 1651
QY 1628 TCCTATGATATATGA 1643
DB 1652 TCCAGTGCCACTCTGA 1667

RESULT 9
AAx89866
ID AAx89866 standard; cdna; 2024 BP.
XX
AC AAx89866;

XX 08-OCT-1999 (first entry)
XX DE L. esculentum germacrene C synthase cDNA (insert pLE 14.2).
XX DE Germacrene C synthase; tomato; transgenic; pathogen resistance;
XX KW flavour alteration; neuroreputical value; ss.
XX OS Lycopersicon esculentum.
XX FH Key Location/Qualifiers
XX FT 32..1678
XX FT CDS
XX FT /*tag= a
XX FT /product= "germacrene C synthase"
XX PN WO9938957-A1.
XX PD 05-AUG-1999.
XX PF 02-FEB-1999; 99WO-US002133.
XX PR 02-FEB-1998; 98US-0073579P.
XX PA (UNIW) UNIV WASHINGTON STATE RES FOUND.
XX PA (REGC) UNIV CALIFORNIA.
XX FI Colby SM, Crock JE, Lemaux PG, Croteau RB;
XX DR WPI; 1999-479180/40.
XX DR P-PSDB; AAY27063.
XX XX
XX PT New Germacrene C synthase protein and nucleic acid useful for producing
XX XX transgenic plants with increased resistance to pathogens.
XX PS Example; Fig 8A-B; 60pp; English.
XX XX
XX CC This represents the nucleotide sequence of the cDNA insert pLE 14.2
XX CC encoding a Lycopersicon esculentum germacrene C synthase protein. The
XX CC germacrene C synthase gene (cDNA insert pLE 20.3) is useful for
XX CC generating transgenic plants with an increased host resistance to
XX CC pathogens including bacteria, rusts, herbivores, microbes and fungi.
XX CC Recombinant nucleic acid comprising a promoter sequence operably to the
XX CC germacrene synthase gene allows tissue specific expression of the protein.
XX CC to deter e.g. leaf-eating herbivores, and allows production of plants
XX CC with an altered flavour and odour profiles to deter e.g. fruit-eating
XX CC herbivores, or to enhance the plants attractiveness for pollinators, or
XX CC increase the neuroreputical value of the plant. Polypeptide fragments of
XX CC the protein are useful as immunogens to raise antibodies, and the
XX CC polynucleotides are useful as probes or primers for detecting, cloning
XX CC and confirming alleles and homologues of the germacrene synthase gene
XX SQ Sequence 204 BP; 705 A; 332 C; 379 G; 608 T; 0 U; 0 Other;
Query Match 21.7%; Score 357.6; DB 2; Length 2024;
Best Local Similarity 52.7%; Pred. No. 1.1e-85;
Matches 861; Conservative 0; Mismatches 749; Indels 24; Gaps 3;
29 TCGCCCCATTGCCAACTTCTCCAGCATTTGGGAGATCAGTTCTCATCTATCAAAA 88
58 TCGCCCCATTGCCAACTTCTCCAGCATTTGGGAGATCAGTTCTCATCTATCAAAA 117
89 GCAAGTAGCAAGGGTGGACACATAGTGAATGATTTAAAAAAGAGTGGCGCACT 148
118 TGAATTAATACTAATCAAGAAAAAGTTGAAGTTGATGATGATCAAGAGACAAATAGAAAAAT 177
149 ACTAAAAAGAGCTTTGGGATATTCCTATGAACATGCCAATTTGTTGAAGCTGATTGATGA 208
178 GCTGGTGGAACTTCGACATAGACACTCAAAA-----GCTTGTGTTGATAGACGC 228
209 AATTCACGCTTGGAAATACCGTATCACTTTGAACGGGAGATGATCATGCAATGCAATG 268
229 GATGCAACGATTTGGGAGTGGCTTATCAATTCGATATGAATGAATGAACATTCATCAAAA 288

QY 269 TATTATGAACAT-----ATGGTGAATACTGGATGGTGGACCGCTCTTCCTT 316
DB 289 CATTTTGTGTCATCGTCCAAACAGAATGATAATGACAACACCTTTAGTGTGTCTCT 348
QY 317 ATGTTCCGCTTTATGCGAAAGCAAGGATATTTATGTTACATGTGATGTTTCAATAAATA 376
DB 349 TCGTTTTCGACTTGTGAGGCAACAGGCCATTAATGATCTTCAGATGTGTTCAAGCAAT 408
QY 377 TAAAGACAAAAATGGAGCGTTTCAAGCAATCGTTAGCTAATGATGTTGAAGTTTGTCTGA 436
DB 409 CACCAACCAAGATGGGAAATTCAGGAAACACATTAATGATGTCCAAGGATTATTGAG 468
QY 437 GTTGTAGAGCAACTTCTATGAGGTACCTGGGAGATTTATATAGAGATGCTCTTGG 496
DB 469 TTTGTATGAAGCATCACATCTGAGAGTGGGTAATGAGGAGATTTCTTGAAGAGCTCTTAC 528
QY 497 TTTTACACGATCTGCTCTAGCATTAATGACAAAAAGATGCTTTTCTACAAACCCGCTCT 556
DB 529 ATTACACCACTCATCTCGAGTCTATTGCTCCAACTTGAGCAATATATACTCTCT 588
QY 557 TTTTACCGAAATACAAACGGGACCTAAAGCAACCCCTTTTGGAAAAAGTTGCCAAGATAGA 616
DB 589 TAAGGTTGAAGTTGGTGAAGCCTTAATCTAGCCTATTGCGATGACTTTTACCAAGGATGG 648
QY 617 GCGGGCGCAGTACATCTCTTCTATCAACAAGATTTCTCATACAAGACTTTTACTTAA 676
DB 649 AGCTAGAAATACATATCCATTTACGAAACAAATGATGACACCACTTTGCTTTTGA 708
QY 677 ACTTGCTAAGTTAGAGTTCAATTTGCTTCACTGTCACAAAGGAAGAGCTCAGGCATGT 736
DB 709 ATTTGCTAATTTGATTTTAACTGCTGCAAAAGTTTCACCAAGAGAGCTTAGTGAICT 768
QY 737 GTCCAAATGGTGAAGCTTTTCGATATCAAGAAAGCAAGCACTTTGTTTGAAGATAGAAT 796
DB 769 TACAAGGTGGTGAAGATTTGGATTTTGAATAAATATCCATATGCAAGAGACAGGTT 828
QY 797 TGTGTAATGCTACTTTTGGGAGCTAGCTTCAGGCTATGAGCCACAGATTTCCGGGCTAG 856
DB 829 GGTGTAGTGTACTCTTGGATATTAGGAGTGTATTTGAGCCAAATAATAGTCGTCGAG 888
QY 857 AGTTTCTTCCAAAAAGCTGTGTTTATAAATCTTTATAGATGACACTTATGATCGCTA 916
DB 889 AAAAATGATGACAAAAGTACTCAACCTGACCTCCATTTATGACGACACTTTTGTATGCTTA 948
QY 917 TGTGCTTATGAAGAACTTAAGATCTTTTACTGAAGCTGTTGAAAGGTGGTCAATTTACATG 976
DB 949 TGCAACTTTTGAAGCACTTGTGACTTTTCAATGATGCAATCCAGAGATGGGATGCTTAATGC 1008
QY 977 CTTAGACACTTCCAGAAATACATGAACCGATATACAAATTTTTCATGGATACATACAC 1036
DB 1009 AATTGATTTCAATACACCATATATGAGACTGCTTTATCAGCTCTTCTAGACATTTACAG 1068
QY 1037 AGAAATGGAGAAATTTCTTCAAAAGGAGGAAACAGATCTATTTTAACTCGCGCAAGA 1096
DB 1069 TGAATGGAAACAGTGTGTTGCCAAAGAGGTAAACTGGACCGTGTATCTATGCAAAAA 1128
QY 1097 ATTTGTGAAGAGTTTGTGTAAGAACCTGATGTTGTAAGCAAAATGGGCAATGA---GGG 1153
DB 1129 TGAGATGAAGAAAGTTGGTGTGAGAGCTATTTTAAAGGAAACCAATGGTTGAATGATGTA 1188
QY 1154 ACACATACCAACCACTGAAGAGCATGATCCAGTTGTAATCATTTACTGCGGTGCTAACT 1213
DB 1189 CCATATCCAAATATAGGAAACAGTGGAGATGCAATCGTAAGTGTGCTATATGAT 1248
QY 1214 GCTTACAACTTGTATCTTGGCATGAGTGAATATTTCAAAAAGAGTCTGTGCAATG 1273
DB 1249 GATATCAACAACTTGTGTTGTCGATATGAAGAAATTTATATCCACGAGACTTTTGAATG 1308
QY 1274 GCGTGTCTGCACTCTCTTTTATAGTACTAGGTATATCTTGGTCAAGCTTAATGA 1333
DB 1309 GTTGATGAATGAGTCTGTGATTTGTCGAGCTTCGCAATGATTTGCCAGAGCAATGAACGA 1368
QY 1334 TCTCATGACCCCAAGCGCCGAGCAAGAAAGAAACATAGTTTTCAGAGCTTTGAAAGTTA 1393

Db 1369 TATGTTGGACATGAAGATGAACAAAGAAAGGAGCATGTAGCTTACATTTATGTAATGTTA 1428
QY 1394 TATGAAGAAATATATGTCATATGAGAGTATGCCCCAACCTTGTATTTTCAAGGAGTAGA 1453
Db 1429 CATGAAGATTATGAGCTTCAAAGCAAGAGACTTACATTAAGTTCCCTGAAAGAGTACAC 1488
QY 1454 AGATGTTGGAAGATATTAACCGAGAGTACTCACAACCTAAACATTAACATTCOAAGCCGTT 1513
Db 1489 CAATGTCATGAAGAGACATAACAAACAATTCCTCCGTCCTCAAGTACCAATGTTTGT 1548
QY 1514 ATTGATGGCTGTGATCTATTGTTGCCAGTTTCTTTGAAAGTTCAATATGTCAGGAAGGATAA 1573
Db 1549 CTTGACAGAGTCTAAATTTGACACGCTGTGGCTGACACGTTATATAAGGAGAAATAC 1608
QY 1574 CTTTACACGATGAGGACAGCAATACAAACATCTCATAAAGTCTTACTCGTTTATCTTAT 1633
Db 1609 ATATTCAACCGCCAAAGGAAACCTTAAACATGATTAATCAATCTACTTAATTGATCTGT 1668
QY 1634 GAGTATATGAGGAT 1647
Db 1669 CAAATATATAATAT 1682

RESULT 10

AAA38935

ID AAA38935 standard; DNA; 2024 BP.

XX AAA38935;

XX 25-AUG-2000 (first entry)

XX Tomato germacrene C synthase DNA sequence SEQ ID NO:51.

XX Synthase; protein co-ordinate data; active site; modification; terpenoid;
KW 3-dimensional coordinate; alpha carbon atom; plant; terpene synthase;
KW isoprenoid; breeding programme; fragrance; flavour; pheromone;
KW defensive agent; pigment; antitumour; steroid hormone;
KW signal transduction pathway; bile acid; affinity purification;
KW photoreceptor; enzymatic synthesis; nutrient supplement;
KW immunological reagent; ds.

XX Lycopersicon esculentum.

XX WO200017327-A2.

XX 30-MAR-2000.

XX 17-SEP-1999; 99WO-US021419.

XX 18-SEP-1998; 98US-0100993P.

XX 22-APR-1999; 99US-0130628P.

XX 23-AUG-1999; 99US-0150262P.

XX (KENT) UNIV KENTUCKY RES DEPT.

XX (SALK) SALK INST BIOLOGICAL STUDIES.

XX Chappell J, Manna KR, Noel JP, Starks CM;

XX WPI; 2000-292839/25.

XX P-PSDB; AAY90856.

XX Novel terpene synthase enzymes, useful for producing terpene

XX enzymes by specific amino acid alterations.

XX Disclosure; Page 429-432; 450pp; English.

XX The present invention describes an isolated terpene synthase (I)
CC comprising a region with at least 20% identity to region 265-535 of a 548
CC amino acid (aa) sequence (Ia), given in AAY90831. (I) contains nine alpha
CC -carbon atoms (alphaC) that have interatomic distances, between each
CC other, within tabulated ranges, have a centre point (within a sphere of

CC radius 2.3 Angstrom) within tabulated ranges, and have an ordered
CC arrangement of R groups (defining as side chains), excluding specific
CC tabulated arrangements (tables given in the specification). (I), and
CC related enzymes, are used to produce a wide range of terpenoids (e.g.
CC cyclic, acyclic, optionally hydroxylated), useful e.g. as fragrances,
CC flavour, pheromones, defensive agents, pigments, antitumour agents,
CC components of signal transduction pathways, precursors of steroid
CC hormones and bile acids, as photoreceptors and as co-factor side chains.
CC Some synthases with little or no catalytic activity (and nucleic acids
CC encoding them) are used as controls in the analysis of products formed by
CC enzymatic synthesis; as nutrient supplements; for affinity purification
CC of isoprenoids; or to develop immunological reagents or nucleic acids for
CC monitoring expression of terpene synthase or inheritance of the gene in
CC plant breeding programs. The new synthases may produce novel terpene
CC products. AAA38910 to AAA38938 and AAY90831 to AAY90859 represent
CC sequences used in the exemplification of the present invention
XX
SQ Sequence 2024 BP; 705 A; 332 C; 379 G; 608 T; 0 U; 0 Other;

Query Match 21.7%; Score 357.6; DB 3; Length 2024;

Best Local Similarity 52.7%; Pred. No. 1.1e-85;

Matches 861; Conservative 0; Mismatches 749; Indels 24; Gaps 3;

QY 29 TCGCCCCATTGCACTTTCTCCAAAGCATTTTCGGGAGATCAGTTTCTCATCTATCAAAA 88
Db 58 TCGCCCCTTGGCTAATTTTCAACCATCTGTTGGGATATCATTTCTTCTTATCTCA 117
QY 89 GCAAGTAGAGCAAGGGGTGGACACAGATAGTGAATGATTTAAAAAGAAAGTCGCGCACT 148
Db 118 TGAATTTACTAATCAAGAAAAGTTTGAAGTTGATGATACAAAGAGACAAATTAGAAAAAT 177
QY 149 ACTAAAAGAGCTTTGGATATTCCTATGAAACATGCCAAATTTGTTGAAGCTGATTGATGA 208
Db 178 GCTGGTGGAACTTGGCAATAGCACTCAAAA-----GCTGTGTTGATAGAGCG 228
QY 209 AATTCACAGCCCTTGAATACCGTATCATTTGACGGGAGATGATCATGCAATTCGAATG 268
Db 229 GATGCAACCAATTTGGAGTGGCTTATCATTTGTAATGAATTAACATCCCAATCAAAA 288
QY 269 TATTTATGAACAT-----ATGGTGATAACTTGGAAATGGTGACCGCTCTTCCTT 316
Db 289 CATTTTGTGATGATCGTCCAAACAGAAATGATATGACACAACTTTACGTTGTGTCTCT 348
QY 317 ATGTTTCCGCTTTATGCGAAAGCAAGGATATTTATGTTATCATGTGATGTTTTCATTA 376
Db 349 TCGTTTTCGACTTGTGAGGCAACAGGCCATTACATGTCTTCAGATGTGTTTCAAGCAAT 408
QY 377 TAAAGACAAAATGGAGCGTTCAGCAATCGTTAGCTTAATGATGTTGAAGTTTGCTTGA 436
Db 409 CACCAACCAAGATGGAAATTCAGGAAACACTTACTAATGATGTCAGAGATTTATTGAG 468
QY 437 GTTGTACGAAGCAACTTCTATGAGGGTACCTGCGGAGATTTATATTAGAAGATGCTCTTGG 496
Db 469 TTTGTTATGAAGCATCACATCTGAGAGTGGCTAATGAGGAGATTTCTTGAAGAGCTCTTAC 528
QY 497 TTTTACAGATCTCGTCTTTAGCAATTAAGCAAAAGATGCTTTTCTTCAAAACCCGCTCT 556
Db 529 ATTTACCACCACTCATCTCGAGTCTATTGTCTCCAACCTTGAGCAATTAATACTCTCT 588
QY 557 TTTTACCGAAATACACGGGCACTAAAGCAACCCCTTTGGAAGGTTGCCAAGAAATAGA 616
Db 589 TAAGGTTGAAGTTGGTGAAGCCCTTAACCTCAGCTTATTCGCATGACTTTACCAAGATGGG 648
QY 617 GCGCGCGCAGTAGTACATTTCCVTTCCTATCAACAAAGATTTCTCATAAACAGACTTTTACTTAA 676
Db 649 AGCTAGAAAATACATATCCATTTACGAAACATGATGCACACACCATTTGCTTTTGA 708
QY 677 ACTTCTAAGTTAGATTCATTTGCTTCAGTCAATTTGCAACAGAAAGCTCAGCCATGT 736
Db 709 ATTTGCTAAATTTGATTTTAAACATGCTGCAAAAGTTTTCACAAAGAGAGCTTAGTGATCT 768
QY 737 GTGCAATGCTGAAGCTTTTCGATATCAAGAGAGCGCACCTTGTTTTAAAGAGATGAAT 796

Db 769 TACAAGGTGGTGAAGATTTGGATTTTCCAAATAAATATCCATATGCAAGACAGAGTT 828
 Qy 797 TGTGAATGCTACTTTTGGGACCTAGGTTAGGCTATGACGACAGTATTTCCGGGCTAG 856
 Db 829 GGTGAGTGTACTTCTGGATATAGGAGTGTATTTGAGCGAAATATAGTCGCGAG 888
 Qy 857 AGTTTCTTCAAAAGCTGTGCTGTATTAATCTCTTATAGATGACACTTATGATGCGTA 916
 Db 889 AAAATGATGACAAAAGTACTCAACCTGACCTCAATTAATGAGACACTTTTATGCTTA 948
 Qy 917 TGGTACTTATGAAGAACTTAAGATCTTTACTGAAGCTGTGAAAGGTGGTCAATTAACATG 976
 Db 949 TGAACCTTTGACGAACCTTGTGACCTTCAATGATGCAATCCAGAGATGGATGCTAATGC 1008
 Qy 977 CTTAGACACACTTCCAGATATCATGAACCGATATACAAATTTATCATGATACATACAC 1036
 Db 1009 AATTGATTCATACAAACCATATATGAGACCTGCTTATCAAGCTCTCTAGACATTTTACAG 1068
 Qy 1037 AGAATGGAAGAAATTTCTTGAAGAGGAGGAAGACAGATCTATTAACTGCGGCAAGA 1096
 Db 1069 TGAATGGAACAGTGTGTCGAAGAGGTAACCTGACCGGTATATCTATGTCGCAAAA 1128
 Qy 1097 ATTTGTGAAGAGTTTGTAGAAACCTGATGTTGAAGCAAAATGGGCAAAATGA---GGG 1153
 Db 1129 TGAGATGAAAAGTTGGTGAGAGCCTATTTTAAGGAACCCCAATGGTTGAATGATGTA 1188
 Qy 1154 ACATATCAACCACTGAAGAGCATGATCCAGTTGTAATCATTTACTGGCGGTGCTAACCT 1213
 Db 1189 CCATATTTCCAAATATAGGAACAAGTGGAGATGCAATCGTAAGTGGTGGCTATATGAT 1248
 Qy 1214 GCTTACAACTCTGTTCTTGGCATGATGATATATTCACAAAGAGCTCTGTGCAATG 1273
 Db 1249 GATATCAACAACTGCTGTTGGTGGTATAGAGAAATTTATATCCACGAGACTTTGAATG 1308
 Qy 1274 GGCTGTCTCTGCACTCTCTTTTATGATATCTAGGATATCTGTTGTCGACGCTAAATGA 1333
 Db 1309 GTTGATGAATGAGTCTGTGATTTGTCGAGCTTCCGCAATGATTCGACAGCAATGAACGA 1368
 Qy 1334 TCTCATGACCCACAGCGCGGACGACGAGAAAGACATGTTTATCGAGCTTGAAGTTA 1393
 Db 1369 TATTGTTGGACATGAAGATGAACAAAGAGAGACATGATGCTTCACTTATTGAATGTTA 1428
 Qy 1394 TATGAAGAAATATATGTCATGAGAGATGATGCCAAACCTTGATTTACAAGGAGTAGA 1453
 Db 1429 CATGAAGATATGAGCTTCAAGCAGAGACTTACATTAAGTCTCTGAAGAGTCTAC 1488
 Qy 1454 AGATGTGGAAAGATATAAACCGAGAGTACCTCAACATAAACAATTTCCAGGCGGTT 1513
 Db 1489 CAATGATGGAAGGACATAAACAACAAATCTCCGTCCAACTGAAGTACCAATGTTGT 1548
 Qy 1514 ATTGATGCTGTGATCTATTGTTGCCAGTTTCTTGAAGTTCATATGCGAGGAGATAA 1573
 Db 1549 CTTGAACGAGTCTTAAATTTGACACGCTGTGCTGACAGCTTATATGAAGAGAAATAC 1608
 Qy 1574 CTTACACGTATGGGAGACGAATACAAACATCTCTATAAAGTCTCTACTGTTTATCCTAT 1633
 Db 1609 ATATTCACCGCAAGGAAACTTAAACATGATTAATCCATTAATGATCTGT 1668
 Qy 1634 GAGTATATGAGAT 1647
 Db 1669 CAAAATATAAAT 1682

RESULT 11
 ID AAA38928

XX AAA38928 standard; DNA; 1944 BP.

AC AAA38928;

XX 25-AUG-2000 (first entry)

DT Potato vetaspiradiene synthase DNA sequence SEQ ID NO:31.

XX

DE

XX

KW Synthase; protein co-ordinate data; active site; modification; terpenoid;
 KW 3-dimensional coordinate; alpha carbon atom; plant; terpene synthase;
 KW isoprenoid; breeding programme; fragrance; flavour; pheromone;
 KW defensive agent; pigment; antitumour; steroid hormone;
 KW signal transduction pathway; bile acid; affinity purification;
 KW photoreceptor; enzymatic synthesis; nutrient supplement;
 KW immunological reagent; ds.
 XX Solanum tuberosum.

XX WO200017327-A2.

XX 30-MAR-2000.

XX 17-SEP-1999; 99WO-US021419.

XX 18-SEP-1998; 98US-0100993P.

XX 22-APR-1999; 99US-0130628P.

XX 23-AUG-1999; 99US-0150262P.

XX (KENT) UNIV KENTUCKY RES DEPT.

XX (SALK) SALK INST BIOLOGICAL STUDIES.

XX Chappell J, Manna KR, Noel JP, Starks CW;

XX WPI; 2000-292839/25.

XX P-PSDB; AAY90843.

XX Novel terpene synthase enzymes, useful for producing terpene
 PT hydrocarbons, e.g. fragrances or antitumor agents, are derived from known
 PT enzymes by specific amino acid alterations.
 XX Disclosure; Page 392-395; 450pp; English.

XX The present invention describes an isolated terpene synthase (I)

CC comprising a region with at least 20% identity to region 265-535 of a 548
 CC amino acid (aa) sequence (Ia), given in AAY90831. (I) contains nine alpha
 CC -carbon atoms (alphaC) that have interatomic distances, between each
 CC other, within tabulated ranges, have a centre point (within a sphere of
 CC radius 2.3 Angstrom) within tabulated ranges, and have an ordered
 CC arrangement of R groups (defining aa side chains), excluding specific
 CC tabulated arrangements (tables given in the specification). (I), and
 CC related enzymes, are used to produce a wide range of terpenoids (e.g.
 CC cyclic, acyclic, optionally hydroxylated), useful e.g. as fragrances,
 CC flavours, pheromones, defensive agents, pigments, antitumour agents,
 CC components of signal transduction pathways, precursors of steroid
 CC hormones and bile acids, as photoreceptors and as co-factor side chains.
 CC Some synthases with little or no catalytic activity (and nucleic acids
 CC encoding them) are used as controls in the analysis of products formed by
 CC enzymatic synthesis; as nutrient supplements; for affinity purification
 CC of isoprenoids; or to develop immunological reagents or nucleic acids for
 CC monitoring expression of terpene synthase or inheritance of the gene in
 CC plant breeding programs. The new synthases may produce novel terpene
 CC products. AAA38910 to AAA38938 and AAY90831 to AAY90859 represent
 CC sequences used in the exemplification of the present invention

XX Sequence 1944 BP; 604 A; 359 C; 398 G; 583 T; 0 U; 0 Other;

Query Match 20.0%; Score 330.4; DB 3; Length 1944;

Best Local Similarity 52.0%; Pred. No. 2.3e-78;

Matches 857; Conservative 0; Mismatches 761; Indels 30; Gaps 4;

QY 14 AGAAGAAAACCTATTTCGCCCATTCGCAACTTCTCCCAAGCATTTGGGGAGATCAGTT 73

Db 92 AGAGGAGGAGATTGTTGCCCCATAGCTGACTTCTCTCCAAAGTCTTTGGGGTATGCTTT 151

QY 74 TCTCATCTATCAAAAGCAAGTAGACGAGGGGTGGAAACAGATAGTGAATGATTTAAAAA 133

Db 152 TCATTCTTCTCCCTCGACATCATGATTCGTGGAAAATATGCTCAAGAGATCGAAACTTT 211

QY 134 AGAAGTGGGCACTACTAAAGAGCTTTGGATATTCCTATGAACATGCCAATTTGTT 193

Db 212 GAGGAACAATCAAGAATTATATTTCTGCTATCTTCTCGAAGAACATTTGGCTGAGAAAT 271

1259 AGAGTCTGTGGAATGGGCTGTCTCTGCACCTCTCTTTTAGATACCTCAGGTATCTTGG 1318
1343 AGATTG--AATGGTTGCCAAGAACCCCTAAAATCTTGGAGGCTAATGTGACGTTATG 1399
1319 TCGACGCTTAATGATCTCATGACCCACAGGCGGACGAGAAAGAAACATAGTTCATC 1378
1400 CCGAGTCATAGATGACATAGCCACCTATGAGTTGAGAAGGTAGAGTCAGATTGCCAC 1459
1379 GAGCCTTGAAGTTATATGAAGGAATAATGTCAATGAGGAGTATGCCCAACCTTGTAT 1438
1460 TGAATGGAATGTTACATGAGAGATTATGGTATCCACAGAAAGGCCATGGAATAATT 1519
1439 TTACAAGGAAGTGAAGATGTGTGAAAGATATAAACCGGAGAGTACCTCACAACCTAAAA 1498
1520 CCAAGAAATGGCTGAGACAGCATGGAAGGATGAAATGAAGGAATCTCTCGACCAATCC 1579
1499 CATTCGAAGCGCTTATTGATGGCTGTGATCTATTGTCGCCAGTTTCTTGAAGTTCAATA 1558
1580 CGTCTCTACAGAGATTCTCTCGCATCTTCAATCTTGTCTCGCATTCGATGTTACTTTA 1639
1559 TGCAGGAA--AGGATAAATTCAACGTATGGGAGACGAATACAAACATCTCATAAAGTC 1615
1640 TAAGCACAAATCAAGATGGATACACTCATCCGGAAGAAAGTACTAAACCTCATATTATTGC 1699
1616 TCTACTCGTTTATCTCTATGAGTATATGA 1643
1700 GTTGTGGTGGACTTATTGAAATTTAA 1727

RESULT 12

AAA38913
ID AAA38913 standard; DNA; 1644 BP.
XX
AC AAA38913;
XX
DT 25-AUG-2000 (first entry)
XX
DE Tobacco 5-epi-aristolochene synthase Y527F DNA SEQ ID NO:7.
KW Synchase; protein co-ordinate data; active site; modification; terpenoid;
KW 3-dimensional coordinate; alpha carbon atom; plant; terpene synthase;
KW isoprenoid; breeding programme; fragrance; flavour; pheromone;
KW defensive agent; pigment; antitumour; steroid hormone;
KW signal transduction pathway; bile acid; affinity purification;
KW photoreceptor; enzymatic synthesis; nutrient supplement;
KW immunological reagent; ds.
XX
OS Nicotiana sp.
OS Synthetic.
XX
PN WO200017327-A2.
XX
PD 30-MAR-2000.
XX
PF 17-SEP-1999; 99WO-US021419.
XX
PR 18-SEP-1998; 98US-0100993P.
PR 22-APR-1999; 99US-0130628P.
PR 23-AUG-1999; 99US-0150262P.
XX
PA (KENT) UNIV KENTUCKY RES DEPT.
PA (SALK) SALK INST BIOLOGICAL STUDIES.
XX
PI Chappell J, Manna KR, Noel JP, Starks CM;
XX
DR WPI; 2000-292839/25.
DR P-PSDB; AAY90834.
XX
PT Novel terpene synthase enzymes, useful for producing terpene
PT hydrocarbons, e.g. fragrances or antitumor agents, are derived from known
PT enzymes by specific amino acid alterations.
XX
PS Example 1; Page 349-352; 450pp; English.

194 GAAGCTGATGATGAAATTCACCGCTTGGATACCGTATCACTTTGAACGGAGATTGA 253
272 GGATCTGATACACATTTGTTGAGCGCTTGGCAATGCTTATCATTTTGAAGAAACAAATAGA 331
254 TCATGCAATTGCAATGATTTATGAAACATATGGT-----GATACTGGAA 298
332 TGATATGTTGGATCAATTTTCAAGAGCAGATCTTAATTTGAGGCTCAGGATACAATGA 391
299 TGGTGACCGCTCTTCTTATGGTTCCGTTCTATGCGAAAGCAAGATATTTATGTTACATG 358
392 TTACAAACITTTATCCGTTCAATTTGCGATTATGAGACACATGTTTACATATCTCCCC 451
359 TGATGTTTCAATTAACATAAAGCAAAATGAGCGGTTCAAGCAATCGTTAGCTAAATGA 418
452 ABAACITTTTATTAGATTTCCAGATGCAAAAGGCAAAATTTAAGAAATCTCTTTGTAACGA 511
419 TGTGAGGTTTGTCTGATTTGTACGAGCAACTTCTATGAGGTTACCTGGGAGATTAT 478
512 CATCAAGGGCTTTTGAACCTTATACGAAGCTCGCATGTAAGGACTCATGGAGAGATAT 571
479 ATTAGAAGATCTCTTGGTTTTACAGCATCTCGTCTTAGCATTAATGACAAAGAGATGCTTT 538
572 TTTGGAAGAGCACTTGTCTTCTACTGCTCATCTTGAATCT-----GCAGCTCC 622
539 TTCTACAAACCCCGCTCTTTTTACGAAATACACGGGCACCTAAAGCAACCCCTTTGGAA 598
623 ACATTTGAAGTCACCTCTGAGTAAGCAAGTACACATGCGCTTGAAGCAATCTCTCCATAA 682
599 AAGTTTGCAGAAATAGAGGGCGGCGAGTACATCTCTTCTATCAACAAAGATCTTCA 658
683 GAGCATTTCCAGAGATTGAGACACGCTACTTCTATCTATCTACGAAAGAGGAGAAACAGAA 742
659 TAACAGACTTTTACTTAACTTTGTAAGTTAGAGTTCAATTTGCTTCACTGATCTGACACAA 718
743 GAATGATGTTGTTCTCAATTTGCAAACTGGACTTCACTTCACTTCACTGATTTGACAA 802
719 GGAAGACTCAGCCATGTGTCAAATGGTGGAAAGCTTTGCGATATCAAGAAGAACGCAAC 778
803 ACAAGAACTTAGTGAAGTATCAAGTGTGCGAAGATTTGGATTTGTGACAACTTCC 862
779 TTGTTTAAGAGATAGAAATTTGATGCTACTTTTGGGACTAGGTTGAGGCTATGAGCC 838
863 ATATGCTAGGATAGACAGTGGAAATGCTACTTTTGGACGATGGGGGTGTATGCTGAACC 922
839 ACAGTATTTCCCGGCTAGAGTTTCTTCAAAAAGCTGTGCTGTATATACTTTATAGA 898
923 TCAATCTCTCAGGCTCGTGTATGCTGCTAGACTATAGCAATGATTTCTATAGTAGA 982
899 TGACACTTATGATGCTATGCTATTAAGAAGCTTAAAGATCTTTTACTGAGCTGTTGA 958
983 TGACACATTCGATGCTTATGCTATGCTCAAGAACTTGAAGATCTACACGATGCCATACA 1042
959 AAGTGTGCTAATTACATGCTTAGACACACTTCCAGAAATACATGAACCGATATACAAAT 1018
1043 GAGGTGGGATATTAGCCAAATTTGATCGGCTCCCTGATTACATGAATAATCAGTTACAAAGC 1102
1019 ATTCAATGATACATACACAGAAATGGAAGAATTTCTGCAAGAGGAGGAAAGAACAGATCT 1078
1103 ACTTTTATGCTTACAAATGATTAATGAATGAGGTTGTCCAAAGATGGTAGATCTGATGT 1162
1079 ATTTAACTGCGGCAAGAAATTTGTAAGAGAGTTTCTTGAAGAACTGATGGTTGAAGCAAA 1138
1163 TGTTCACTACGCAAGAAAGAAAGAAATCGTGAGAACTATTTGTGGAAGCAAA 1222
1139 ATGGGCAATGAGGAGACATACCAACCACTGAAGAGCATGATCCAGTTGTAATCATTTAC 1198
1223 ATGTTTCATTGAAGGATATATGCGCCAGTCTCTGAGTATCTTAGCAATGCTAGTAC 1282
1199 TGGGGGTGCTAACTGCTTACAAACCTTGTATCTTGGCATGAGTATATTACACAA 1258
1283 CAGCACTTATTACTTGTCTTACACTACATCTTATTTGGGCATGAAGTCTGCTAACAGCA 1342

XX The present invention describes an isolated terpene synthase (I) comprising a region with at least 20% identity to region 285-535 of a 548 amino acid (aa) sequence (Ia), given in AAY90831. (I) contains nine alpha-carbon atoms (alphaC) that have interatomic distances, between each other, within tabulated ranges, have a centre point (within a sphere of radius 2.3 Angstrom) within tabulated ranges, and have an ordered arrangement of R groups (defining a side chain), excluding specific tabulated arrangements (tables given in the specification). (I), and related enzymes, are used to produce a wide range of terpenoids (e.g. cyclic, acyclic, optionally hydroxylated), useful e.g. as fragrances, flavours, pheromones, defensive agents, pigments, anticancer agents, components of signal transduction pathways, precursors of steroid hormones and bile acids, as photoreceptors and as co-factor side chains. Some synthases with little or no catalytic activity (and nucleic acids encoding them) are used as controls in the analysis of products formed by enzymatic synthesis; as nutrient supplements; for affinity purification of isoprenoids; or to develop immunological reagents or nucleic acids for monitoring expression of terpene synthase or inheritance of the gene in plant breeding programs. The new synthases may produce novel terpene products. AAY38910 to AAY38938 and AAY90831 to AAY90859 represent sequences used in the exemplification of the present invention

SQ Sequence 1644 BP; 517 A; 322 C; 332 G; 473 T; 0 U; 0 Other;

Query Match 19.2%; Score 316.6; DB 3; Length 1644;
 Best Local Similarity 51.9%; Pred. No. 1.1e-74;
 Matches 846; Conservative 0; Mismatches 764; Indels 21; Gaps 5;

QY 15 GAAGAAAACCTATTGCGCCCATTCGCCAATCTCTCCAGCAATTTGGGAGATCAGTTT 74
 DB 28 GAAGAAGAGATTGTTGCGCCCGTGGCGAGCTTCTCCCTAGTCTCTGGGTGATCAGTTC 87

QY 75 CTCATCTATCAAAAGCAAGTAGAGCAAGGGGTGGAAACAGATAGTGAATGATTTAAAAAAA 134
 DB 88 CTTTCATCTCCA---TTGATATCAGGTTGCGGAAAGTATGCTCAAGAGATTGAGCA 144

QY 135 GAAGTGGCGCACTACTAAAGAGAGCTTTGGATATTCCTATGAAACATGCGCAATTTGTTG 194
 DB 145 TTGAAGGAACAAACGAGGAGTATGCTGTAGCAACCGGAAGGAAATTTGGCCGATACATTG 204

QY 195 AAGCTGATTGATGAATTCACCGCTTGGAAATACCGTATCATCTTGAACGGGAGATTGAT 254
 DB 205 AATTGATTGATGATTGGAAGCGCTTGGTATATCTCCACCTTTGAGAAGAAATTTGAT 264

QY 255 CATGATTGCAATGTATTATGAAACATATGTTGATAAAGTGAATGGTGAACCGCTCTTCC 314
 DB 265 GAGATTTGGATCAGATTTACAAACCAAACTCAAATGCAATGTTGTCCACCTCTGCA 324

QY 315 TTATGGTTCGCTTATGCGAAGCAAGGATATTTATGATGATGTTTCAATAAC 374
 DB 325 CTTCAATTTGATTTGCTCAGGCAACACCGGTTTCAACATCTCTCCTGAAATTTTCAGCAA 384

QY 375 TATAAGACAAATAATGGAGCGTTCAAGCAATGTTAGCTAATGATTTGAGGTTTGCCTT 434
 DB 385 TTCCAGATGAAATTTGGCAATTTCAAGGAGTCTCTTGTAGTATGATCTCTTAGGATTTATTA 444

QY 435 GAGTTGTACGAAGCAACTTCTATGAGGGTACCTGGGGAGATTTATATAGAAGATGCTCTT 494
 DB 445 AACTTGTATGAGCTTCACATGTAAGACTCATGCTGACGATATCTTAGAGAGCACTT 504

QY 495 GGTTTTACAGATCTGCTTTAGCATTTATGACAAAGATGCTTTTCTCAAAACCCCGCT 554
 DB 505 GCTTTTCTC-----CACTATCCATCTTTGAATCTGCAGCTCCACATTTGAAATCTCCA 555

QY 555 CTTTTTACCGAATAACAGGCACTAAAGCAACCCCTTTGGAAAGGTTGCCAAGATA 614
 DB 556 CTTAGGAGCAAGTACATGCCCTTTGAGCAATGTTTGCACAGGGGTGTTCTTAGAGTC 615

QY 615 GAGGGGGCGAGTACATTCCTT---TCTATCAACAACAAGATTCTCATCAACAAGACTTTA 671
 DB 616 GAGACCCGATTCTTCATCTCATCAATCTATGACAGGAAACATCGAAGATATATGTTA 675

QY 672 CTTAAACCTTGCTAAGTTAGAGTTCAATTTGCTTCAGTCAATTCGCAACAAGAGAGCTCAGC 731
 DB 676 CTTTCGATTTGCCAAATTTGGATTTCAACTTTGCTCCAGATGTTGCAACAAGAGACTTGCT 735

QY 732 CATGTGTCAAATGGTGGAAAGCTTTTCGATATCAAGAAAGACGACCTTGTGTTAAGAGAT 791
 DB 736 CAAGTATCAAGGTGGTGGAAAGATTTGGATTTGTAACAACACTTCCATATGCTAGAGAT 795

QY 792 AGAATTTTGAATCTACTTTTGGGGACTAGGTTTCAGGCTATGAGCCACAGATTTCCCGG 851
 DB 796 CGAGTAGTGAATGCTACTTTTGGGCATTAGAGGTTTATTTTGGCCTCAATACTCTCAA 855

QY 852 GCTAGAGTTTCTTCACAAAAGCTTGTGCTGTTTATACTCTTATAGATGACACTTATGAT 911
 DB 856 GCTCGCTGATGCTGCTGTTAGACCATATCAATGATTTTCGATTTGCTGATGACACTTTGAT 915

QY 912 GCGTATGCTACTTATGAAGAACTTTAAGATCTTTACTGAAGCTGTTTGAAGGTTGGTCAAT 971
 DB 916 GCTTACGGTACAGTTAAGAACTTTGAGGCATACACAGATGCCATACAAGATGGGATATC 975

QY 972 ACATGCTTAGACACACTTCCAGAAATACATGAACCGATATACAAATTTATTCATGGATACA 1031
 DB 976 AACGAAATTTGATCGGCTTCTGATTACATGAAATCACTTATAAAGCTATTCTAGATCTC 1035

QY 1032 TACACAAATGGAAGAAATTTCTTGCAAAAGGAGGAAACAGATCTATTTAACCTCGGC 1091
 DB 1036 TACAAGGATTATGAAGAAGAAATTTCTAGTCCGGAAGATCTCATATTTGTGCTCCATGCA 1095

QY 1092 AAGAAATTTGGAAGAGTTTGTAGAAACCTGTGTTGAAGCAAAATTTGGCAAAATGAG 1151
 DB 1096 ATAGAAAGATGAAGAAGTAGTAGAAATTTAATGTGCGAGTCAACATGGTTTATTGAA 1155

QY 1152 GGACACATACCAACCACTGAAGAGCATGATCCAGTTGTAATCATCTGCGCGTGTAAAC 1211
 DB 1156 GGATATATGCCACCTGTTCTGAATACCTAAGCAATGCACTAGCAACTACCACATATTATC 1215

QY 1212 CTGCTTACAACAACTTGTATCTTGGCATGATGATATTTCAAAAGAGAGCTCTCGAA 1271
 DB 1216 TACCTCGCAACATCGTATTTGGGCGATGA---AGTCTGCCACGGAGCAAGATTTTGAG 1272

QY 1272 TGGGCTGTCTCTGCACCTCTCTTTTAGATACCTCAGGTACTTGGTTCGACGCTTAAAT 1331
 DB 1273 TGGTTGTCAAGATCCAAAATTTCTTGAAGCTAGTGTAAATTTATGTCGAGTTATCGAT 1332

QY 1332 GATCTCATGCCCAAGCGCGCAAGAAAGAAACATAGTTTCATCGAGCCTTGAAAGT 1391
 DB 1333 GACACGCCACGTACGAGGTTTGAGAAAGCAGGGGACAAATTTGCAACTGGAATTGAGTGC 1392

QY 1392 TATATGAGGATATATATGTCATGAGGAGTATGCCCAAACTTGTGATTTACAGGAGATA 1451
 DB 1393 TGATGAGAGATTATGTTATATCAACAAAAGAGGCAATGGCTAAATTTCAAATATGGCT 1452

QY 1452 GAAGATCTGTGGAAGATATAAACCGAGAGTACCTCAACAACTAAAACATTTCCAAAGCGC 1511
 DB 1453 GAGACACATGGAAGAATATTAAGAAGACTTCTTAGGCCACTCCCGTCTCTACAGAA 1512

QY 1512 TTATTTGATGCTGTGATCTATTTTGTGCCAGTTTCTTGAAGTTCAATATGCAAGAAA---G 1568
 DB 1513 TTTTAACTCTCTTCTCAATCTTCTGCTGCTGATTTGTTGAGGTTACATATATACAACTTA 1572

QY 1569 GATTAATTTACAGTATGGGAGAGCAATACAAACATCTCAATAAGTCTCTACTCTGTTAT 1628
 DB 1573 GATGGATTTCACTATCCGGAGAAAGTCTTAAACCTCACTTATTAACCTACTTGTGGAC 1632

QY 1629 CCTATGAGTAT 1639
 DB 1633 TCCATCAAAAT 1643

RESULT 13
 AAA38915
 ID AAA38915 standard; DNA; 1644 BP.
 XX

AAA38915;
25-AUG-2000 (first entry)
Tobacco 5-epi-aristolochene synthase Y406X/L407X DNA SEQ ID NO:11.
Synthase; protein co-ordinate data; active site; modification; terpenoid;
3-dimensional coordinate; alpha carbon atom; plant; terpene synthase;
isoprenoid; breeding programme; fragrance; flavour; pheromone;
defensive agent; pigment; antitumour; steroid hormone;
signal transduction pathway; bile acid; affinity purification;
photoreceptor; enzymatic synthesis; nutrient supplement;
immunological reagent; ds.
Nicotiana sp.
Synthetic.
WO200017327-A2.
30-MAR-2000.
17-SEP-1999; 99WO-US021419.
18-SEP-1998; 98US-0100993P.
22-APR-1999; 99US-0130628P.
23-AUG-1999; 99US-0150262P.
(KENT) UNIV KENTUCKY RES DEPT.
(SALK) SALK INST BIOLOGICAL STUDIES.
Chappell J, Manna KR, Noel JP, Starks CM;
WPI; 2000-292839/25.
P-PSDB; AAY90836.
Novel terpene synthase enzymes, useful for producing terpene
hydrocarbons, e.g. fragrances or antitumor agents, are derived from known
enzymes by specific amino acid alterations.
Example 1; Page 358-360; 450pp; English.
The present invention describes an isolated terpene synthase (1)
comprising a region with at least 20% identity to region 285-535 of a 548
amino acid (aa) sequence (1a), given in AAY90831. (1) contains nine alpha
-carbon atoms (alphaC) that have interatomic distances, between each
other, within tabulated ranges, have a centre point (within a sphere of
radius 2.3 Angstrom) within tabulated ranges, and have an ordered
arrangement of R groups (defining aa side chains), excluding specific
tabulated arrangements (tables given in the specification). (1), and
related enzymes, are used to produce a wide range of terpenoids (e.g.
cyclic, acyclic, optionally hydroxylated), useful e.g. as fragrances,
flavours, pheromones, defensive agents, pigments, antitumour agents,
components of signal transduction pathways, precursors of steroid
hormones and bile acids, as photoreceptors and as co-factor side chains.
encoding them) are used as controls in the analysis of products formed by
enzymatic synthesis; as nutrient supplements; for affinity purification
of isoprenoids; or to develop immunological reagents or nucleic acids for
monitoring expression of terpene synthase or inheritance of the gene in
plant breeding programs. The new synthases may produce novel terpene
products. AAA38910 to AAA38938 and AAY90831 to AAY90859 represent
sequences used in the exemplification of the present invention
Sequence 1644 BP; 517 A; 319 C; 332 G; 470 T; 0 U; 6 Other;
Query Match 19.1%; Score 315.4; DB 3; Length 1644;
Best Local Similarity 51.7%; Pred. No. 2.4e-74;
Matches 843; Conservative 1; Mismatches 766; Indels 21; Gaps 5;
15 GAAGAAAAACCTATTGCGCCCATTCGCAACTTTCCTCCCAAGCATTTGCGGAGATCAGTTT 74
28 GAAGAGAGATGTTGCGCCCGTCGCGGACTTTCCTCCTAGTCTCTGGGTGATCAGTTT 87

QY 75 CTCATCTATCAAAAGCAAGTAGAGCAAGGGGTGGAACAGATAGTAGATGATTTAAAAAAA 134
DB 88 CTTTCATTCTCCA---TTGATATCAGGTTGCGGAAAAGTATGCTCAAGAGATTGAAGCA 144
QY 135 GAAGTGGCGCAACTACTATAAAGAGCTTTGGATATTCCTATATGAAAACATGCCAATTGTTG 194
DB 145 TTGAAGGAACAAACGAGGAGTATGCTTTAGCAACCGGAAGGAAATTTGGCGGATACATTG 204
QY 195 AAGCTGATTGATGAATTCACCGCTTTGGAATACCGTATCATCTTGAACGGGAGATTGAT 254
DB 205 AATTTGATTGACATTATTGAAACGCTTTGGTATATCTCCACCTTTGAGAAGAAATTTGAT 264
QY 255 CATGCATTGCAATGATTATTGAACAATATGGTGAATACTGGAATGGTGACCGCTCTTCC 314
DB 265 GAGATTTTGGATCAGATTTTACAACCAAAACTCAAACTGCAATGATTTGTGCACCTCTGCA 324
QY 315 TTATGTTTCCGCTTTATGCGAAAGCAAGATATATTGTTACATGATGATGTTTCAATAAC 374
DB 325 CTTCAATTTGATGCTCAGGCAACACCGGTTTCAACATCTCTCCTGAAATTTTTCAGCAA 384
QY 375 TATAAAGACAAAATGGAGCGTTTCAAGCAATCGTTAGCTAATGATGTTGAAGGTTTGCCT 434
DB 385 TTCCAAGATGAATGGCAAAATTCAGGAGTCTCTTGTAGTATGATGTTTGAAGTATTA 444
QY 435 GAGTTGTACGAAGCAACTTCTATGAGGTTACCTGGGAGATATATATAGAAGATGCTCTT 494
DB 445 AACTTGTATGAAGCTTCACATGTAAGGACTCATGCTGACGATATCTTTAGAAGACGCACTT 504
QY 495 GGTTTTACAGATCTCGCTTTAGCATTTATGACAAAAGATGCTTTTCTACAAACCCGCT 554
DB 505 GCTTTTCTC-----CACTATCATCTTGAATCTGCAGCTCCACATTTTGAATCTTCCA 555
QY 555 CTTTTCACGAAATCAACGGGCACCTTAAAGCAACCCCTTTGGAAAAGGTTGCCAAGAATA 614
DB 556 CTTAGGAGCAAGTACACATGCTTTCAGGAGTCTCTTGTAGTATGATGTTTTCAGAGTC 615
QY 615 GAGCGGCGCAGTACATTCCTT---TCTATCAACAAAGATTTCTCATACAAGACTTTA 671
DB 616 GAGACCCGATTTCTCATCTCATCAATCTATGACAAAGAAACAAATCGAAGAAATAATGTGA 675
QY 672 CTTAAACTTGTCAAGTTAGATTCAATTTGCTTCACTTCACTTGCACAAAGGAGAGTCAGC 731
DB 676 CTTGATTTGCCAAATTTGGATTTCAACTTGTCTCCAGATGTTGCACAAACAAAGAACTTGT 735
QY 732 CATGTGTGCAATGTGTGAAGCTTTGATATCAAGAAAGCAACGACCTTGTTTAAAGAGAT 791
DB 736 CAAGTATCAAGGTGTGTGAAGATTTGGATTTGTAAACAACTTCCATATGCTAGAT 795
QY 792 AGAATTGTTGAATGCTACTTTTGGGAGTGTGTTGAGGCTATGAGCCACAGTATTCGCGG 851
DB 796 CGAGTAGTTGAATGCTACTTTTGGGCAATTAGGAGTTTATTTTGGAGCTCAATCTCTCAA 855
QY 852 GCTAGAGTTTCTTCAAAAGCTGTGCTGTATTAACCTCTTATAGATGACATTTATGAT 911
DB 856 GCTCGCGTCTGCTCGTTTGAAGCAATATCAATGATTTTCGATTTGTCGATGACACCTTTGAT 915
QY 912 CGGTATGTTACTTATGAAGAACTTAAAGATCTTTTCTGGAAGCTGTTGAAAAGGTGCTCAAT 971
DB 916 CTTACGTTACAGTTTAAAGAACTTGAAGCATACACAGATGCCATACAAGATGGGATATC 975
QY 972 ACATGCTTTAGACACACTTCCAGAAATACATGAACCGGATATACAAATTTATTCATGGATACA 1031
DB 976 AACGAAATTTGATCGGCTTCTGATTACATGAAATTCAGTTATAAAGCTATTTCTAGATCTC 1035
QY 1032 TACACAGAAATGGAAGATTTCTTCAAGAGGAGGAGAAACAGATCTATTATTAATCGCGGC 1091
DB 1036 TACAAGGATTATGAAGAGAAATTTGTAGTGCAGGAGATCTCATATTTGTCTGCCATGCA 1095
QY 1092 AAAGAAATTTGGAAGAGCTTTGTTAGAACCTGTGTTGAAGCAAAATTTGGGCAATGAG 1151
DB 1096 ATAGAAAGATGAAGAAAGTATGAGAATTAATATGTCGAGTCAACATGGTTTATTGAA 1155
QY 1152 GGACACATACCACCACTTGAAGAGCATGATCCAGTTTGTAAATCATTAATCTGCGGTGCTAAC 1211

Db 1156 GGATATATGCCACCTGTTCTCTGAATACCTAAGCAATGCACCTAGCACTACCAATATAC 1215
Qy 1212 CTGCTTACAACTGTTATCTTGGCATGAGTATATATTCACAAAGAGTCTGCGAA 1271
Db 1216 NNSNNSGCGAACAATCGTATTTTGGCATGA---AGTCTGCCACGAGCAAGATTTTGAG 1272
Qy 1272 TGGGCTGTCTCTGCACCTCTCTCTTTTATAGATACCTCAGGTATATCTTGGTGCAGCCCTAAAT 1331
Db 1273 TGGTGTCTCAAGAATCCAAAATCTTGNAGCTAGTGTAAATATATGTGAGTATCGAT 1332
Qy 1332 GATCTATGACCCCAAGCGGAGCAAGAAAGAAACATAGTTCATGCGACCTTGAAGT 1391
Db 1333 GACACAGCCAGTACGAGGTTGAGAAAAGCAGGGGACAAATTCGCAACTGGAATTTGAGTC 1392
Qy 1392 TATATGAGGAATATAATGTCAATGAGGAGTATGCCAAACCTTGATTTTACAGGAGTA 1451
Db 1393 TGCATGAGATTTTGGTATATCAAAAAGGCAATAGGCTTAATTTCAAAATATGGCT 1452
Qy 1452 GAAGATGTGTGGAAGATATAAACGAGAGTACCTCACAACCTTAAAAAACATTCGAAGCGG 1511
Db 1453 GAGACAGCATGGAAGATATTAATGAAGGACTTCTTAGGCCACCTCCGCTCTACAGAA 1512
Qy 1512 TTATGTAGGCTGATCTATTTTGGCAGTTTCTTGAAGTTCAATATGACAGAA---G 1568
Db 1513 TTTTAACTCTCTATTTCTCAATCTTCTCGTATGTTGAGGTTTACATATACACAATCTA 1572
Qy 1569 GATAACTTTCACAGTATGCGGAGAGCAATACAAACATCTCATAAAGTCTCTACTCTTTAT 1628
Db 1573 GATGATACATCTATCCGAGAAAGCTTAAACCTCAATATTAACCTACTTGTGGAC 1632
Qy 1629 CCTATGAGTAT 1639
Db 1633 TCCATCAAAAT 1643

RESULT 14
AAA38910
ID AAA38910 standard; DNA; 1644 BP.
XX
AC AAA38910;
XX
DT 25-AUG-2000 (first entry)
DE Tobacco 5-epi-aristolochene synthase DNA SEQ ID NO:1.
XX
KW Synthase; protein co-ordinate data; active site; modification; terpenoid;
KW 3-dimensional coordinate; alpha carbon atom; plant; terpene synthase;
KW isoprenoid; breeding programme; fragrance; flavour; pheromone;
KW defensive agent; pigment; antitumour; steroid hormone;
KW signal transduction pathway; bile acid; affinity purification;
KW photoreceptor; enzymatic synthesis; nutrient supplement;
XX immunological reagent; ds.
OS Nicotiana sp.
XX
PN WO200017327-A2.
XX
PD 30-MAR-2000.
XX
PF 17-SEP-1999; 99WO-US021419.
XX
PR 18-SEP-1998; 98US-0100993P.
PR 22-APR-1999; 99US-0130628P.
PR 23-AUG-1999; 99US-0150262P.
XX
XX (KENT) UNIV KENTUCKY RES DEPT.
PA (SALK) SALK INST BIOLOGICAL STUDIES.
XX
XX Chappell J, Manna KR, Noel JP, Starks CM;
PI WPI; 2000-292839/25.
XX P-PSDB; AAY90831.
DR

XX Novel terpene synthase enzymes, useful for producing terpene
PT hydrocarbons, e.g. fragrances or antitumor agents, are derived from known
PT enzymes by specific amino acid alterations.
XX
XX Disclosure; Page 336-339; 45Opp; English.
XX
CC The present invention describes an isolated terpene synthase (I)
CC comprising a region with at least 20% identity to region 265-535 of a 548
CC amino acid (aa) sequence (Ia), given in AAY90831. (I) contains nine alpha
CC -carbon atoms (alphaC) that have interatomic distances, between each
CC other, within tabulated ranges, have a centre point (within a sphere of
CC radius 2.3 Angstrom) within tabulated ranges, and have an ordered
CC arrangement of R groups (defining a side chain), excluding specific
CC tabulated arrangements (tables given in the specification). (I), and
CC related enzymes, are used to produce a wide range of terpenoids (e.g.
CC cyclic, acyclic, optionally hydroxylated), useful e.g. as fragrances,
CC flavour, pheromones, defensive agents, pigments, antitumour agents,
CC components of signal transduction pathways, precursors of steroid
CC hormones and bile acids, as photoreceptors and as co-factor side chains.
CC Some synthases with little or no catalytic activity (and nucleic acids
CC encoding them) are used as controls in the analysis of products formed by
CC enzymatic synthesis; as nutrient supplements; for affinity purification
CC of isoprenoids; or to develop immunological reagents or nucleic acids for
CC monitoring expression of terpene synthase or inheritance of the gene in
CC plant breeding programs. The new synthases may produce novel terpene
CC products. AAA38910 to AAA38938 and AAY90831 to AAY90859 represent
CC sequences used in the exemplification of the present invention
XX
SQ Sequence 1644 BP; 518 A; 322 C; 332 G; 472 T; 0 U; 0 Other;
Query Match 19.1%; Score 315; DB 3; Length 1644;
Best Local Similarity 51.8%; Pred. No. 3e-74;
Matches 845; Conservative 0; Mismatches 765; Indels 21; Gaps 5;
Qy 15 GAAGAAAAACCTATTCCGCCCATTTGCCAACTTTCTTCAAGCATTTGGGGAGATCAAGTTT 74
Db 28 GAAGAGAGATTTTCCGCCCGTCCGCACTTCTCCCTAGTCTCTGGGGTATCAGTTC 87
Qy 75 CTCATCTATCAAGCAAGTAGAGCAAGGGTGGACAGATAGTGAATGATTTAAAAAA 134
Db 88 CTTTCACTTCCCA---TTGATATATCAGGTTGCGGAAAAGTATGCTCAAGAGATTGAACA 144
Qy 135 GAAATGCGGCACTACTATAAAGAGCTTTGGATATTCTTATGAAACATGCCAAATTTGTTG 194
Db 145 TTGAGGACAAACAGGAGATGCTGTTAGCAACCGGAGGAATTTGGCGATACATTG 204
Qy 195 AAGCTGATTGATGAATCAAGCCCTTGGAAATACCGTATCACTTTGAACGGGAGATTGAT 254
Db 205 AATTTGATTGACATTTATTGAACGCTTTGGTATATCTTACCACCTTGAGAAAGAAATTCAT 264
Qy 255 CATGCAATTCGAATGATTTTATGAAACATATGGTGATACTGGATGGTACCCTCTTCC 314
Db 265 GAGATTTTGGATCAGATTTTACAAACCAAACTCAAACTGCAATGATTTGTGCACCTCTGCA 324
Qy 315 TTATGGTTCCTGCTTATGCGAAGCAAGGATATATGTTTACATGTGATGTTTCAATAAC 374
Db 325 CTTCAATTTGATTTGCTCAGGCAACAGGTTTCAACATCTCTCCTGAAATTTTCAAGCAA 384
Qy 375 TATAAGACAAAAATGGAGCGTTCAAGCAATCGTTAGCTAATGATGTTGAAGTTTGCTT 434
Db 385 TTCCAAAGATGAAATTTGCAAAATTTCAAGGAGTCTCTTGTCTAGTGTCTTAGGATTATTA 444
Qy 435 GAGTTGTACGAGCAACTTCTATGAGGTACCTGGGAGATTATATTAGAGATGCTCTT 494
Db 445 AACTTGTATGAAGCTTTCACATGTAGGACTCATGCTGACGATATCTTTAGAGAGCGACTT 504
Qy 495 GOTTTTTACAGTCTGCTTGTAGCATTTATGACAAAAGATGCTTTTCTCAAAACCCGCT 554
Db 505 GCTTTCTC-----CACTATCCATCTTGTATCTGAGTCTCCACATTTGAAATCTCCA 555
Qy 555 CTTTTTACGAAATACAAACGGGCACTAAAGCAACCCCTTTGGAAAAGGTTGCCAAGATA 614

Db 556 CTTAGGGAGCAAGTGACACATGCGCTTGAGCAATGTTTGACAAAGGCGTGTCTAGAGTC 615
Qy 615 GAGGGCGGCGAGTACATTCCTT---TCTATCAACAACAAGATTCCTATACAGACTTTA 671
Db 616 GAGACCGGATTCCTCATCTCAATCTATGACAGGAACAATCGAAGATAATGTGTTA 675
Qy 672 CTTAACTTGTCAAGTGAAGTGAAGTTCCTTCACTTATGACAGGAAGAGCTCAGC 731
Db 676 CTTGATTTGCCAAATGCAATGCAATTCCTTCACTTATGACAGGAAGAGCTCAGC 735
Qy 732 CATGTGTGCAAAATGTTGGAAGAGCTTCGATATCAAGAAGACCCACTTGTGTTAAGAT 791
Db 736 CAAGTATCAAGGTGTGGAAGAGTTCGATATCAAGAAGACCCACTTGTGTTAAGAT 795
Qy 792 AGAATGTTGATGCTACTTTTGGGAGTTCAGGCTATGAGCACACAGTATCCCGG 851
Db 796 CGAGTAGTGAATGCTACTTTTGGGAGTTCAGGCTATGAGGCTCAATCTCTCAA 855
Qy 852 GCTAGAGTTTCTTCAAAAAGCTGTGCTGTATTAACCTTATAGATGACACTTATGAT 911
Db 856 GCTGCGTCATGCTCGTTAAGACCATATCAATGATTTGCGATGACACCTTTGAT 915
Qy 912 GCGTATGCTACTTATGAAGACTTAAGATCTTCTGTAAGCTGTTGAAAGGTGGTCAAT 971
Db 916 GCTTACGGTACAGTTAAAGAACTTGAGGCATACACAGATGCCATACAAAGATGGGATC 975
Qy 972 ACATGCTTAGACACACTTCCAGATACATGAACCGATATACAAATTTATTCATGGATACA 1031
Db 976 AACGAAATGATCGGCTTCTGATACATGAATTCAGTTATTAAGCTATCTAGATCTC 1035
Qy 1032 TACACAGAAATGGAAGATTTCTTCAAGAGGGGAGAACACAGATCTATTTAACTCGGGC 1091
Db 1036 TACAAGGATTAAGAAAGAAATGCTAGTCCGGAAGATCTCATATTGCTGCCATGCA 1095
Qy 1092 AAAGAAATTTGAAGAGTTTGTGTAAGACCTGTGTTGAAGCAAAATGGGCAATGAG 1151
Db 1096 ATAGAAGAAATGAAGAAAGTATGAAGAAATTAATGTCGAGTCAACATGGTTATTGAA 1155
Qy 1152 GGACATACCAACCACTGAAGAGATGATCCAGTTGTAATCATATTGCGGCGTCTAAC 1211
Db 1156 GGATATATGCCACCTGTTCTGAAATACCTAAGCAATGCTAGCAATCCACATATTAC 1215
Qy 1212 CTGCTTACAAACAATGTTATCTTGGCATGAGTGATATATTCACAAAGAGTCTGCGAA 1271
Db 1216 TACCTCGGCAACAACATGCTATTTGGCATGA--AGTCTGCCAGCGGCAAGATTTGAG 1272
Qy 1272 TGGGCTGTCTGCACTCTCTTTTAGATACCTAGGTATCTAGGTATCTGCGCGCTAAAT 1331
Db 1273 TGGTTGTCAAGAAATPCCAAAATTTCTGAAGCTAGTGTAAATATATGTCGAGTTATCGAT 1332
Qy 1332 CATCTCATGACCCACAGCGCAGCAAGAAAGAAACATAGTTTCATCGAGCCTTGAAGT 1391
Db 1333 GACACAGCCAGTACAGGTTGAGAAAGCAGGGGACAAATTCGCAACTGGAAATTGAGTGC 1392
Qy 1392 TATATGAAGGAATATATGTCAATGAGGAGTATGCCAAACCTTGATTTTACAGGAAGTA 1451
Db 1393 TGCATGAGAGATTATGGTATATCAACAAGAGGCAATGGCTAAATTTCAAATATGGCT 1452
Qy 1452 GAAGATGTGGAAGATATAACCGAGAGTACCTCAACATTAACAAATTCAGAGCGG 1511
Db 1453 GAGACAGATGGAAGATATTAATGAAGACTTCTTAGGCCACTCCCTCTCTACAGAA 1512
Qy 1512 TTATTGAGGCTGTGATCTATTGTCAGTTTCTTGAAGTTCAATATGAGGAA--G 1568
Db 1513 TTTTAACTCCTATTTCTCAATCTTCTGCTGATTTGAGGTTACATATATACAAATCTA 1572
Qy 1569 GATAACTTCAACGATGAGGAGCAATACAAACATCTCATAAAGTCTCTACTCGTTTAT 1628
Db 1573 GATGATACACTCATCCGAGAAAGTCTTAAACCTCAATATTAACTTCTTGTGGAC 1632
Qy 1629 CCTATGAGTAT 1639
Db 1633 TCCATCAAAAT 1643

RESULT 15

AAA38912
ID AAA38912 standard; DNA; 1644 BP.

XX AC AAA38912;

XX DT 25-AUG-2000 (first entry)

XX DE Tobacco 5-epi-aristolochene synthase Y520F DNA SEQ ID NO:5.

XX KW Synthase; protein co-ordinate data; active site; modification; terpenoid;
3-dimensional coordinate; alpha carbon atom; plant; terpene synthase;
isoprenoid; breeding programme; fragrance; flavour; pheromone;
defensive agent; pigment; antitumour; steroid hormone;
signal transduction pathway; bile acid; affinity purification;
photoreceptor; enzymatic synthesis; nutrient supplement;
immunological reagent; ds.

XX OS Nicotiana sp.

XX OS Synthetic.

XX XX WO200017327-A2.

XX XX 30-MAR-2000.

XX XX 17-SEP-1999; 99WO-US021419.

XX PR 18-SEP-1998; 98US-0100993P.

XX PR 22-APR-1999; 99US-0130628P.

XX PR 23-AUG-1999; 99US-0150262P.

XX PA (KENT) UNIV KENTUCKY RES DEPT.

XX PA (SALK) SALK INST BIOLOGICAL STUDIES.

XX PI Chappell J, Manna KR, Noel JP, Starks CM;

XX DR WPI; 2000-292839/25.

XX DR P-PSDB; AAY90833.

XX PT Novel terpene synthase enzymes, useful for producing terpene
hydrocarbons, e.g. fragrances or antitumor agents, are derived from known
enzymes by specific amino acid alterations.

XX PS Example 1; Page 345-347; 450pp; English.

XX CC The present invention describes an isolated terpene synthase (I)
comprising a region with at least 20% identity to region 265-535 of a 548
amino acid (aa) sequence (Ia), given in AAY90831. (I) contains nine alpha
-carbon atoms (alphaC) that have interatomic distances, between each
other, within tabulated ranges, have a centre point (within a sphere of
radius 2.3 Angstrom) within tabulated ranges, and have an ordered
arrangement of R groups (defining aa side chains), excluding specific
tabulated arrangements (tables given in the specification). (I), and
related enzymes, are used to produce a wide range of terpenoids (e.g.
cyclic, acyclic, optionally hydroxylated), useful e.g. as fragrances,
flavours, pheromones, defensive agents, pigments, antitumour agents,
components of signal transduction pathways, precursors of steroid
hormones and bile acids, as photoreceptors and as co-factor side chains.
Some synthases with little or no catalytic activity (and nucleic acids
encoding them) are used as controls in the analysis of products formed by
enzymatic synthesis; as nutrient supplements; for affinity purification
of isoprenoids; or to develop immunological reagents or nucleic acids for
monitoring expression of terpene synthase or inheritance of the gene in
plant breeding programs. The new synthases may produce novel terpene
products. AAA38910 to AAA38938 and AAY90831 to AAY90859 represent
sequences used in the exemplification of the present invention

XX SQ Sequence 1644 BP; 517 A; 323 C; 332 G; 472 T; 0 U; 0 Other;

Query Match

Best Local Similarity 18.9%; Score 311.8; DB 3; Length 1644;

Pred. No. 2.2e-73;

Matches 843; Conservative 0; Mismatches 767; Indels 21; Gaps 5;			
Qy	15	GAAGAAAACCTATTTCGCCCAATTGCAACTTTCCTCCCAAGCAATTTGGGAGATCAGTTT	74
Db	28	GAAGAAAGAAATTTGTCGCCCTTCGCGCACTTCTCCCTAGTCTCTGGGGTGATCAGTTC	87
Qy	75	CTCATCTATCAAAAGCAAGTAGAGCAAGGGGTGGAACAGATAGTGAATGATTTAAAAAA	134
Db	88	CTTTCAATCTCA---TTGATTAATCAGGTTCGGAAAAGTATGCTCAAGAGATTGAACA	144
Qy	135	GAAGTGGCGGAACCTACTATAAAGAGCTTTGGATATTCCTATGAACAATGCCAATTTGTTG	194
Db	145	TTGAAGGAACAAACGAGGAGTATGCTTTAGCAACCGGAAGGAAATTTGGCCGATACATTG	204
Qy	195	AAGCTGATTGATGAATCAACGCTTGAATACGTCATCTTCACTTCAACCGGAGATTGAT	254
Db	205	AATTTGATTGACATTTAGAACGCTTTGGTATATCTTACCACTTTGAGAAAGAAATTTGAT	264
Qy	255	CATGCAATTCGAATGATTTATTAAGAAATATGTTGATTAACCTGGAATGCTGACCCCTCTTC	314
Db	265	GAGATTTTGGATCAGATTTACAAACCAAACTCAAACCTGCAATGTTTGTGCACCTCTGCA	324
Qy	315	TTATGGTTCGCTTATGCGAAGCAAGGATATTAATGTTACATGATGATTTTCAATAAC	374
Db	325	CTTCAATTTGATTTGCTCAGGCAACACCGTTTCAACATCTCTCTGAAATTTTTCAGCAA	384
Qy	375	TATTAAGCAAAAATGGAGCGTTCAAGCAATCGTTAGCTAATGATCTTGAAGTTTGCCT	434
Db	385	TTCCAAGATGAATGGCAATTCAGGAGTCTCTTCTAGTATGATCTCTTAGGATTTATTA	444
Qy	435	GAGTTGTACGAACCACTTCTATGAGGTACCTGGGGAGATATATTAAGAAGATGCTCT	494
Db	445	AACTTGTATGAAGCTTCACATGTAAGGACTCATGCTGACGATATCTTAGAGACGACCT	504
Qy	495	GGTTTACAGATCTGCTTTAGCAATTAAGCAAAAGATGCTTTTCTCAAAACCCGCT	554
Db	505	GCTTTCTC-----CACTATCATCTTGAATCTGCAGCTCCACATTTGAAATCTCCA	555
Qy	555	CTTTTTCAGGAATCAACGGGCATTAAGCAACCCCTTTTGGAAAGAGTTGCCAAGATA	614
Db	556	CTTAGGAGCAAGTACACATGCCCTTTAGCAATGTTTCCACAAAGGTGTTCTTAGATC	615
Qy	615	GAGCGGCGCAGTACATTCCTT---TCTATCAACAACAAGATTTCTATCAACAAGCTTTA	671
Db	616	GAGACCCGATTTCTCATCTCATCAATCTATGACAGGAACAATCGAAGAATAATGTTGTTA	675
Qy	672	CTTAACCTTGCTAAGTTAGATTTCAATTTGCTTCACTCATTTGCAACAGGAAGCTCAGC	731
Db	676	CTTCGATTTGCCAAATGGATTTCACTTTGCTCCAGATGTTGCAACAACAAGAACTTGCT	735
Qy	732	CATGTTGCCAATGTTGGAAGCTTTCCGATATCAAGAAAGCAACGACCTTTGTTTAAAGAT	791
Db	736	CAAGTATCAAGTGTGGAAGATTTGGATTTTGTAAACAACCTTCCATATGCTAGAT	795
Qy	792	AGAATTTGTTGAATGCTACTTTTGGGACTAGGTTCAAGCTATGAGCCACAGTATTCGCCG	851
Db	796	CGAGTAGTTGAATGCTACTTTTGGCCATTAGAGTTTATTTTGGCCTCAATACITCTCAA	855
Qy	852	GCTAGAGTTTCTTCAAAAAGCTGTTGCTGTTTAACTCTTATAGATGACACTTATGAT	911
Db	856	GCTCGCTCATGCTCGTTTAAAGACCATATCAATGATTTTCGATTTGATGACACCTTTGAT	915
Qy	912	GCGTATGCTACTTATGAAGAACTTAAGATCTTTTACTGAGCTGTTGAAAGGTGGTCAAT	971
Db	916	GCTTACGGTACAGTTTAAAGAACTTAGGCAATACACAGATGCCATCAAGATGGGATATC	975
Qy	972	ACATGCTTTAGACACATTTCCAGATACATGAACCGATATACAAATTTATTCATGGATACA	1031
Db	976	AACGAAATTTGATCGCTTCTCTGATACATGAATAATCAGTTTAAAGCTATTTCTAGATCTC	1035
Qy	1032	TACACAGAAATGGAAGAAATTTCTTCCAAAGGAGGAGAACAGATCTATTTAACTCGCGC	1091
Db	1036	TACAAGGATTATGAAGGAATTTGTCTAGTGGCGGAAGATCTCATATTTGCTGCCATGCA	1095

Search completed: June 7, 2004, 09:34:50
Job time : 744 secs

Qy	1092	AAAGATTTTGTGAAGAGTTTGTAGAAACCTGATGTTGAAGCAAAATGGCAAAATGAG	1151
Db	1096	ATAGAAAAGAAATGAAGAAAGTAGTAAGAAATTAATATGCGAGTCAACATGGTTTATTGAA	1155
Qy	1152	GGACACATACCAACCACTGAAGAGCATGATCCAGTTGTAATCATTTACTTGGCGGTGCTAAC	1211
Db	1156	GGATATATGCCACTGTTTCTTGAATACTTAAGCAATGCTAGCAACTACACATATTAC	1215
Qy	1212	CTGCTTACACAACCTTGTATCTTTGGCATAGTATGATTAATTCACAAAGAGTCTGTGCA	1271
Db	1216	TACCTCGCAACACATCGTATTTGGGCATGA---AGTCTGCCACGAGCAAGATTTTGAG	1272
Qy	1272	TGGGCTGCTCTGCACTCTCTTTTGTAGATACTCAGGTATACTTGGTTCGACGCTAAAT	1331
Db	1273	TGGTGTCAAGAAATCCAAAATTTCTTGAAGCTAGTGAATTTATGTCGAGTTTTCGAT	1332
Qy	1332	GATCTCATGACCCACAAAGCGCGGCAAGAAAGAAAACATAGTTTCATCGAGCCTTGAAAGT	1391
Db	1333	GACACGCCACGTACGAGGTTTGAGAAAAGCAGGGGACAAAATTGCAACTGGAATTTGAGTGC	1392
Qy	1392	TATATGAAGAAATATAATGTCAATGAGGAGTATGCCCAAACTTGTGATTTACAGGAAGTA	1451
Db	1393	TGCATGAGAGATTATGGTATATCAACAAAAGAGGCAATGGCTAAATTTCAAAATATGGCT	1452
Qy	1452	GAAGATGTGTGGAAGATATAAACCGAGAGTACCTCAACAACATAAAACATTTCCAAAGCCG	1511
Db	1453	GAGACAGCATGGAAAGATATTAATGAAGGACTTCTTAGGCCCACTCCCGTCTTACAGAA	1512
Qy	1512	TTATGATGCTGTGATCTATTTTGTGCCAGTTTCTTGAAGTTTCAATATGCGAGAAA---G	1568
Db	1513	TTTTTAACTCTTATCTCAATCTTGTCTGATTTGTGAGGTTTACATTCACACAATCTA	1572
Qy	1569	GATAACTTTCACAGTATGGGAGAGCAATACAAACATCTCAATAAGTCTCTACTCGTTTAT	1628
Db	1573	GATGGATACACTCATCCGAGAAAGTCTTAAACCTCACATTTAATTAACCTTCTTGGGAC	1632
Qy	1629	CCTATGAGTAT	1639
Db	1633	TCCATCAAAAT	1643

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2004, 09:06:39 ; Search time 151 Seconds
(without alignments)
6060.358 Million cell updates/sec

Title: US-09-763-822A-13
Perfect score: 1649
Sequence: 1 ccattggcacttacagaagaa.....ctatgagtatataggatccc 1649

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:**
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:**
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:**
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:**
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:**
5: /cgn2_6/ptodata/2/ina/PCUTS_COMB.seq:**
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:**

ALIGNMENTS

RESULT 1
US-09-601-091-1
; Sequence 1, Application US/09601091
; Patent No. 6342380
; GENERAL INFORMATION:
; APPLICANT: Colby, S. M. et al.
; TITLE OF INVENTION: Germacrene C Synthase Gene of Lycopersicon Esculentum
; FILE REFERENCE: 4630-55758
; CURRENT APPLICATION NUMBER: US/09/601,091
; CURRENT FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: PCT/US99/02133
; PRIOR FILING DATE: 1999-02-02
; PRIOR APPLICATION NUMBER: US 60/073,579
; PRIOR FILING DATE: 1998-02-02
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1879
; TYPE: DNA
; ORGANISM: Lycopersicon esculentum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (39)..(1685)
US-09-601-091-1

Query Match 22.1%; Score 364; DB 4; Length 1879;
Best Local Similarity 52.9%; Pred. No. 1.7e-97;
Matches 865; Conservative 0; Mismatches 745; Indels 24; Gaps 3;

QY	29	TCGCCCATTCGCAACTTTCCTCCAGCATTGGGGAGATCAGTTTCTCATCTATCAAAA	88
DB	65	TCGCCCATTCGCTTAATTTTCCACCCTCTGTTGGGATATCAATTCCTTCTTACTCA	124
QY	89	GCAAGTAGAGCAGAGGGTGGACAGATAGTGAATTTTAAAAAAGAGTGGGCAACT	148
DB	125	TGAAATTAATCAAGAAAAAGTTGAAGTTGATGATACAAAGACAAATAGAAAT	184
QY	149	ACTAAAAGAGCTTTGGATATTCCTATGAACATGCCAAATTTGTTGAAGCTGATTGATGA	208
DB	185	GCTGTGGAAACTTCGACAAATAGCACTCAAAA-----GCTTGTGTTGATAGACGC	235
QY	209	AATTCACCGCTTGGATACCGTATCTTTCGACGGGAGATTGATCATGATTCGAATG	268
DB	236	GATGCAACGATTGGGAGTGGCTTATCTATCGATAATGAAATGAAACATCAATCAAAA	295
QY	269	TATTTATGAACAT-----ATGCTGATAACTGGAATGGTGCCTCTCTCTT	316
DB	296	CATTTTGTGATCGTCCAAACAGATGATGATGACAAACCTTACGTGTGCTCT	355
QY	317	ATGTTTCCGCTTTATGCGAAAGCAAGGATATATGTTACATGTGATGTTTCAATACTA	376

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	364	22.1	1879	4	US-09-601-091-1 Sequence 1, Appli
2	357.6	21.7	2024	4	US-09-601-091-3 Sequence 3, Appli
3	357.6	21.7	2024	4	US-09-398-395A-5 Sequence 51, Appl
4	357.6	21.7	2024	4	US-09-887-586A-51 Sequence 51, Appl
5	357.6	21.7	2024	4	US-09-895-752-51 Sequence 51, Appl
6	357.6	21.7	2024	4	US-09-903-012B-51 Sequence 51, Appl
7	357.6	21.7	2024	4	US-09-900-797-51 Sequence 51, Appl
8	330.4	20.0	1944	4	US-09-398-395A-31 Sequence 31, Appl
9	330.4	20.0	1944	4	US-09-887-586A-31 Sequence 31, Appl
10	330.4	20.0	1944	4	US-09-895-752-31 Sequence 31, Appl
11	330.4	20.0	1944	4	US-09-903-012B-31 Sequence 31, Appl
12	330.4	20.0	1944	4	US-09-900-797-31 Sequence 31, Appl
13	320.6	19.4	1671	4	US-09-398-395A-1 Sequence 1, Appli
14	320.6	19.4	1671	4	US-09-887-586A-1 Sequence 1, Appli
15	320.6	19.4	1671	4	US-09-895-752-1 Sequence 1, Appli
16	320.6	19.4	1671	4	US-09-903-012B-1 Sequence 1, Appli
17	320.6	19.4	1671	4	US-09-900-797-1 Sequence 1, Appli
18	316.6	19.2	1644	4	US-09-398-395A-7 Sequence 7, Appli
19	316.6	19.2	1644	4	US-09-887-586A-7 Sequence 7, Appli
20	316.6	19.2	1644	4	US-09-895-752-7 Sequence 7, Appli
21	316.6	19.2	1644	4	US-09-903-012B-7 Sequence 7, Appli
22	316.6	19.2	1644	4	US-09-900-797-7 Sequence 7, Appli
23	315.4	19.1	1644	4	US-09-398-395A-11 Sequence 11, Appl
24	315.4	19.1	1644	4	US-09-887-586A-11 Sequence 11, Appl
25	315.4	19.1	1644	4	US-09-895-752-11 Sequence 11, Appl
26	315.4	19.1	1644	4	US-09-903-012B-11 Sequence 11, Appl
27	315.4	19.1	1644	4	US-09-900-797-11 Sequence 11, Appl

Db 356 TCGTTTTCGACTGTGTGAGCAACAAGGCCATTACATGTCTTCAGATGTGTTCAGCAATT 415
QY 377 TAAAGACAAAATGAGCGTTCAGCAATCGTTAGCTAATGATGTGTGAAGGTTGCTTGA 436
Db 416 CACCACCAAGATGGGAAATTCAGGAAACACTTACTAATGATGTCCAAAGGATATTGAG 475
QY 437 GTTGTACGAAGCAACTTCTATCAGGGTACTCTGGGAGATTATATTAGAAGATGCTCTTGG 496
Db 476 TTTGTATGAAGCATCACATCTGAGAGTGGTATGAGGAGATTCCTTGAAGAGCTCTTAC 535
QY 497 TTTTACAGATCTGCTTTAGCATTTATGACAAAGATGCTTTTCTACAAACCCCGCTCT 556
Db 536 ATTTACCACCACTCATCTCGAGTCTATTCTCTCAACTTGGAGCAATAATAAATACCTCT 595
QY 557 TTTTACCGAATFACAAGGGCACTAAAGCAACCCCTTTGGAAGAGTTCGCAAGATAGA 616
Db 596 TAAGTTGAAGTTGGTGAAGCCTTAACCTCAGCCTATTGCGATGACTTTACCAAGGATGG 655
QY 617 GCGCGCGAGTACATTCCTTTCTATCAACAAGAGATTCTCATACAAGACTTTTACTTAA 676
Db 656 AGCTAGAAAATACATATCCATTTACGAAACATGATGACACCACTTTGCTTTGAA 715
QY 677 ACTTCTAAGTTAGAGTTCAATTTGCTTCAGTCAITTCGCAAGGAGAGCTCAGCCATGT 736
Db 716 ATTTGCTAAATTTGATTTTAAATGCTGCAAAAGTTTCACCAAGAGAGCTTAGTGATCT 775
QY 737 GTGCAATGCTGGAAGCTTTGATATCAAGAGAGCGACCTGTTTAAGAGATAGAT 796
Db 776 TACAAGGTGGTGAAGATTTGGATTTGCAATAAATATCCATATGCAAGAGACAGGT 835
QY 797 TGTGAATCTACTTTTGGGGAGTGTTCAGGCTATGAGCCACAGTATTCGCGGCTAG 856
Db 836 GGTGTAGTTTACTTCTGATATTAGAGTGTATTTGAGCCAAATATAGTCTGCGAG 895
QY 857 AGTTTCTTCAAAAAGCTGTGCTGTTTAACTCTTTATAGATGACACTTTATGATGCTTA 916
Db 896 AAAAATGATGACAAAAGTACTCAACCTGACCTCCATTATTGAGGACACTTTTGTATGCTTA 955
QY 917 TGGTACTTATGAGAACTTAAGTCTTTACTGAGGCTGTTGAAGGTGGTCAATTACATG 976
Db 956 TGCAACCTTTGACGAACCTTGTGACTTTCAATGATCAATCCAGAGATGGATGCTAATGC 1015
QY 977 CTTAGACACTTCCAGATAACATGAAACCGATATACAAATTTATCATGATACATACAC 1036
Db 1016 AATTGATTCAAATACACCATATATAGAGCTGCTTATCAGCTCTTACAGATTTACAG 1075
QY 1037 AGAAATGGAAGAAATTTCTTGAAGGAGGAGAAAGAGATCTATTATTAACCTGCGGCAAGA 1096
Db 1076 TGAATGGAACAAAGTGTGTCGCAAGAGGTAACCTGGACCGGTATATCTATGCAAAAAA 1135
QY 1097 ATTTGTGAAGAGTTTGTGTAGAAACCTGATGTTGAAGCAAAATGGCCRAATGA--GGG 1153
Db 1136 TGAGATGAAAAGTTGGTGAGAGCCCTATTATTAAGGAAACCCATGGTGAATGATGTGA 1195
QY 1154 ACACATACCACCACTGAAGAGCATGATCCAGTTGTAATCATTTACTGGCGGTGCTAACCT 1213
Db 1196 CCATATTCAAAATATGAGGAACAAGTGGAGATGCAATCGTAAGTCTGCTATATGAT 1255
QY 1214 GTTTACAAACAATTGTTTGGCATGAGTGATATATTACAAAAGAGTCTGTGCAATG 1273
Db 1256 GATATCAACAACCTGCTTGGTGGTATAGAAGAAATTTATATCCACGAGACTTTTGAATG 1315
QY 1274 GCGTGTCTGACCTCTCTTTTATAGTACTCAGGTATATCTGGTGGAGCGCTTAATGA 1333
Db 1316 GTTGATGAATGAGTGTGTGATTTGTCAGCTTCGCAATGATTGCCAGAGCAATGAACGA 1375
QY 1334 TCTCATGACCCCAAGGCGAGCAAGAAAGAAACATAGTTTCATCGAGCCTTGAAAGTTA 1393
Db 1376 TATTGTTGACATGAAGATGAAACAGAAAGAGACATGTAGTCTCACTTATTGAATGTTA 1435
QY 1394 TATGAAGGAATAATAATGTCAATGAGGAGTATGCCAAACCTTGATTTACAGGAAGTAGA 1453

Db 1436 CATGAAGAATTATGGAGCTTCAAAGCAAGAGACTTACATTAAAGTTCTCTGAAAGAGGTCA 1495
QY 1454 AGATGTGTGAAGATATAACCGAGAGTACCTCACAACTAAAAACATTCCCAAGCGGTT 1513
Db 1496 CAATGATGAAGAGACATAAACAACAACTTCTCGTCAACTGAAGTACCAATGTTTGT 1555
QY 1514 ATTGATGGCTGTGATCTATTGTTGCCAGTTTCTTTGAAAGTTCATATGCGAGAAAGATAA 1573
Db 1556 CCTTGAACGAGTCTTAAATTTGACACGTTGGCTGACACGTTATATAAAGAGAAAGATAC 1615
QY 1574 CTTCAACAGTATGGGAGACGATACAAACATCTCATAAAGTCTCTACTCGTTTATCCTAT 1633
Db 1616 ATATACAAACGCCAAAGGAAACCTTAAAAACATGATTAACTCAATTAATGAAATCTGT 1675
QY 1634 GAGTATATGAGGAT 1647
Db 1676 CAAAATATAATAT 1689

RESULT 2
US-09-601-091-3
; Sequence 3, Application US/09601091
; Patent No. 8342380
; GENERAL INFORMATION:
; APPLICANT: Colby, S. M. et al.
; TITLE OF INVENTION: Germacrene C Synthase Gene of Lycopersicon Esculentum
; FILE REFERENCE: 4630-55758
; CURRENT APPLICATION NUMBER: US/09/601,091
; PRIOR APPLICATION NUMBER: PCT/US99/02133
; PRIOR FILING DATE: 1999-02-02
; PRIOR APPLICATION NUMBER: US 60/073,579
; PRIOR FILING DATE: 1998-02-02
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2024
; TYPE: DNA
; ORGANISM: Lycopersicon esculentum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (32)..(1678)
US-09-601-091-3

Query Match 21.7%; Score 357.6; DB 4; Length 2024;
Best Local Similarity 52.7%; Pred. No. 1.4e-95;
Matches 861; Conservative 0; Mismatches 749; Indels 24; Gaps 3;

QY 29 TCGCCCATTCGCCAACTTCTCCCAAGCATTTGGGAGATCAGTTTCTCATCTATCAAAA 89
Db 58 TCGCCCTTGGCTAATTTTCACCCATCTGTTGGGATATCATTTCTTCTTACTCA 117
QY 89 GCAAGTAGACAGGGGTGGACAGATAGTGAATGATTTAAAAAAGAGTGGCGCACT 148
Db 118 TGAATTTACTAATCAAGAAAAAGTTGAAGTTGATGATACAAAGAGACAATTAGAAAAAT 177
QY 149 ACTAAAAGAACTTTGGATATTCTTATGAAACATGCCAATTTGTTGAACTGATGATGA 208
Db 178 GCTGGTGAACCTTCGCAATAGCACTCAAAA-----GCTTGTGTGTAGACGC 228
QY 209 AATTCAACGCTTGGATACCGTATCACTTTGAACGGGAGATGATCATGCAATGCAATG 268
Db 229 GATGCAACGATTTGGAGTGGCTTATCAATTGATTAATGAAATTTGAAACATCCATTCAAAA 288
QY 269 TATTATGAACAT-----ATGGTGAATACCTGGAATGGTGGCGCTCTTCCCTT 316
Db 289 CATTTTGTATGCTGCTCCAAACAGAAATGATTAATGACAAACCTTTAGCTGTGTCTCT 348
QY 317 ATGTTTCCGTCTTATGCGAAAGCAAGGATATTATGTTACATGTGATGTTTTCATAACTA 376
Db 349 TCGTTTTCGACTTGTGGGCAACAGGCCATTACATGCTTCAGATGTGTTCAGCAATT 408
QY 377 TAAAGACAAAAATGGAGCGTTCAAGCAATCGTTAGCTAATGATGTTGAAGGTTTGTCTGA 436

Db 409 CACCAACCAAGATGGAAATTCACAGAAACACTTACTAATGATGTCCAAGGATTATTGAG 468
QY 437 GTTGTACCAAGCAACTTCTATGAGGAGTACTCGGAGATATATTAGAGAGTCTCTTGG 496
Db 469 TTTGTATGAGCATCACATCTGAGAGTGGTAATGAGGAGATCTTTGAAGAAGCTCTTAC 528
QY 497 TTTTACAGCATCTCGTCTTAGCATATGACAAAAGATGCTTTTCTACAAAACCCGCTCT 556
Db 529 ATTTACCACCACTCATCTCGAGTCTATTGTCTCCAACTTGAGCAATAATAAATCTCT 588
QY 557 TTTTACCGAATACAAACGGGCACTAAAGCAACCCCTTTGGAAAAGGTTGCCAAGATAGA 616
Db 589 TAAGGTGAAGTGGTGAAGCCCTTAACCTCAGCTATTTCGCATGACTTTACCAAGGATGG 648
QY 617 GCGGCGCAGTACATCTCTTTCTATCAACAAAGATTTCTCATACAAAGACTTTTACTTAA 676
Db 649 AGCTAGAAAATACATATCAATTTAGAAAACAAATGATGACACCACTTTGCTTTTGA 708
QY 677 ACTTGCTAAGTTAGAGTTCAATTTGCTTCAGTCAITGCAAGGAGAGCTCAGCCATGT 736
Db 709 ATTTGCTAAATGGATTTTAAACATGCTGCAAAAGTTTACCAAAGAGAGCTTAGTGATCT 768
QY 737 GTGCAAAATGGTGGAAAGCTTTTCGATATCAAGAAAGCAACCTTGTGTTAAGAGATAGAAT 796
Db 769 TACAAGGTGGTGGAAAGATTTGGATTTTGCAAAATAAATATCCATATGCAAGAGACAGTT 828
QY 797 TGTGTAATGCTACTTTTCGGGACTAGGTTTCAGGCTATGAGCCACAGTATTCGGGCTAG 856
Db 829 GGTGAGTGTACTTCTCGATATTAGGAGTGTATTTTGAGCCAAAATATAGTCGTCGAG 888
QY 857 AGTTTTCTTCAAAAGCTGTGCTTTATATCTTATAGATGACACTTATGATCGTA 916
Db 889 AAAAATGATGACAAAGTACTCAACTGACCTCCATTTATGACGACACTTTTGTGCTTA 948
QY 917 TGGTACTTATGAAGAACTTATAGATCTTACTGAAAGCTGTGAAAGGTGTCATTAATG 976
Db 949 TGCAACCTTTGACGAACTGTGACTTTCAATGATGCAATCCAGAGATGGATGCTAATGC 1008
QY 977 CTTAGACACTTCCAGATACATAAACCAGATATACAAATTTATCATGATACATACAC 1036
Db 1009 AATTGATTCATACAAACATATATAGAGACCTGCTTATCAAGCTCTCTTAGACATTTACAG 1068
QY 1037 AGAAATGGAAGAAATTTCTGCAAGAGGAGGAGACAGATCTATTAACTGGGCAAGA 1096
Db 1069 TGAATGGAACAGTGTGTCCAAAGAGAGTTAACTGGACCGTGATATATATGCAAAA 1128
QY 1097 ATTTGTGAAGAGTTGTTGTAAGAACTGATGTTGAAGCAAAATGGCAAAATGA---GGG 1153
Db 1129 TGAGATGAAAAAGTTGTTGAGAGCCTATTTTAAAGGAAACCCCAATGGTTGAATGTGTA 1188
QY 1154 ACATATACCAACCACTGAAGAGCATGATCCAGTTGTAATCATTTACTGGCGGTGCTAACT 1213
Db 1189 CCAATTTCCAAAATATAGGAAACAGTGGAGAAATGCAATCGTAAGTGTGCTGCTATATGAT 1248
QY 1214 GCTTACAACAACTGTGTTATCTTGGCATGAGTGATATATTACAAAAGAGTCTGTGCAATG 1273
Db 1249 GATATCAACAACTTGTGTCGGTATAGAGAAATTTATATCCACAGAGACTTTTGAATG 1308
QY 1274 GGCTGTCTCTGCACCTCTCTTTTAGATACTCAGGTATATCTTGTGCGACGCTAAATGA 1333
Db 1309 GTTGATGAATGAGTCTGTGATTGTTTCGAGCTTCCGCAATTTGATTCGACAGCAATGAACA 1368
QY 1334 TCTCATACCCACAGGCGGAGCAAGAAACATATAGTTTCATGAGCTTTGAAAGTTA 1393
Db 1369 TATTGTTGGACATGAAGATGAACAAAGAGGACATAGCTTCACTTATTGAATGTTA 1428
QY 1394 TATGAAGGAATATAATGTCATAGGAGTATGCCCAACCTTGATTTTCAAGGAAGTAGA 1453
Db 1429 CATGAAGATATTGGAGCTTCAAGCAAGAGACTTTACATTAAGTTCTCGAAGAGGTCAC 1488
QY 1454 AGATGTGGAAGATATAAACCGAGAGTACCTCAACACTAAAAACATTTCCAAGGCCGTT 1513

Db 1489 CAATGCATGGAAGACATATAAAACAAATCTCCGCTCCAACTGAAGTACCAATGTTTGT 1548
QY 1514 ATTGATGGCTGTGATCTATTGTCGCCAGTTTCTTGAAGTTCAATATGAGGAAAGATAA 1573
Db 1549 CCTTGAACGAGTTCTAAATTTGACACGCTGTGCGTGCACGTTATATAAGGAGAAAGATAC 1608
QY 1574 CTTTCAACGATATGGAGACGAATACAAACATCTCATAAAGTCTCTACTCGTTTATCCTAT 1633
Db 1609 ATATTCAACGCCCAAGGAAAACCTTAAACATGATTAATCCAATCTAATTAATCTGT 1668
QY 1634 GAGTATATGAGGAT 1647
Db 1669 CAAAATATAAATAT 1682

RESULT 3
US-09-398-395A-51
; Sequence 51, Application US/09398395A
; Patent No. 6468772
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 64687721, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/398,395A
; CURRENT FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 51
; LENGTH: 2024
; TYPE: DNA
; ORGANISM: Lycopersicon esculentum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (32)...(1675)
; OTHER INFORMATION: VFNT germacrene C synthase
US-09-398-395A-51

Query Match 21.7%; Score 357.6; DB 4; Length 2024;
Best Local Similarity 52.7%; Pred. No. 1.4e-95;
Matches 861; Conservative 0; Mismatches 749; Indels 24; Gaps 3;

QY 29 TCGCCCATTCGCAACTTTCCTCCAGCATTTGGGAGATCAGTTTCTCATCTATCAAAA 88
Db 58 TCGCCCATTCGCTAAATTTTCCACCATCTGTTGGGATATCATTTCTTCTTACTCA 117
QY 89 GCAAGTAGAGCAAGGGGTGGAACAGATAGTGAATGATTTAAAAAAGAGTGGCAACT 148
Db 118 TGAATTAATAATCAAGAAAAGTTGAGTTGATGAGTACAAAGAGACAAATAGAAAAT 177
QY 149 ACTAAAGAGCTTTGGATATTCCTATGAACATGCCAATTTGTTGAAGCTGATTGATGA 208
Db 178 GCTGTGTAAGAACTTGGACAATAGCACTCAAAA-----GCTGTGTTGATAGACGC 228
QY 209 AATTCAACGCTTGGAAATACCGTATCACTTTGAACGGAGATTGATCATGATTCGAATG 268
Db 229 GATCAACGATTTGGAGTGGCTTATCATTTTCGATTAATGAAATTTGAAACATCCATTCAAAA 288
QY 269 TATTTATGAACAT-----ATGTTGATTAATGGAATGGTACCGCTCTTCTT 316
Db 289 CATTTTGTGATCTGTCCTCAACAGAAATGATATGACAAACCTTTACGTTGTGCTCT 348
QY 317 ATGTTTCCGCTTTATGCGAAAGCAAGGATATTGTTACATGTGATGTTTTCAATACTA 376
Db 349 TCGTTTTCGACTTGTGAGGCAACAGGCCATTACATGCTTTCAGATGTGTTCAAGCAATT 408

QY 377 TAAAGCAAAATGGAGCGTTCCAGCAATCGTTAGCTAATGATGATGAGGTTTGGTTGA 436
Db 409 CACCACCAAGATGGGAATTCAGGAACACCTTACTAATGATGTTCCAAAGGATTATTGAG 468
QY 437 GTTGTACGAAGCAACTTCTATGAGGCTACCTGGGAGATTATATTAGAAGATGCTCTGG 496
Db 469 TTGTGATGAGCATCACAATCGAGAGTGGTAATGAGGAGATCTTGAAGAGCTCTTAC 528
QY 497 TTTTACACATCTGCTCTTATGATGATGATGATGATGATGATGATGATGATGATGATG 556
Db 529 ATTTACCAACCT 588
QY 557 TTTTACCGAAATACACAGCGGCTACTAAGCAACCCCTTTGGAAAGGTTGCCAAGATAGA 616
Db 589 TAAGGTTGAAGTTGGTGAAGCTTAACTCAGCTATTTCGATGACTTTTACCAAGGATGG 648
QY 617 GGGGGCGAGTACATTCCTTTCTATCAACACAGATTTCTATTAACAGACTTTTACTTAA 676
Db 649 AGCTAGAAATACATATCCATTTACGAAACCAATGATGACACCCCACTTTCTTTTGA 708
QY 677 ACTTGTAAGTTAGAGTTCAATTTGCTTCAAGTCAATGCAAGGAGAGCTCAGCCATGT 736
Db 709 ATTGCTAATTTGAATTTTAACTGCTGCAAGTTTACCAAGAGAGCTTAGTGATCT 768
QY 737 GTGCAAAATGGTGAAGAGCTTTCGATATCAAGAGAGAGAGAGAGAGAGAGAGAGAG 796
Db 769 TACAAGTGGTGAAGAGATTTGGAATTTTGAATTTTGAATTTTGAATTTTGAATTTGA 828
QY 797 TGTGAATGCTACTTTTGGAGAGTGGTTCAGCTATGAGCCACAGTATTCCTCCGGGCTAG 856
Db 829 GGTGAGTGTACTTCTGATATTAGAGAGTGTATTTGAGCCCAATATAGTCTGGGAG 888
QY 857 AGTTTCTTCAAAAAGCTTGTGCTTTAATACTCTTATAGATGACACTTATGATGCTGA 916
Db 889 AAAAAATGATGACAAAAGTACTCAACCTGACCTCCATTTTGAAGAGAGAGAGAGAG 948
QY 917 TGTGCTTATGAGAGCTTAAAGATCTTACTGAGCTGTGTAAGGTCGTCATTTACATG 976
Db 949 TGCACCTTTGAGAGACTTGTGACTTTCAATGATGCAATCCAGAGATGGATGCTAATGC 1008
QY 977 CTTAGACACTTCCAGATATACATGAACCGATATACAAATTTATTCATGATACATACAC 1036
Db 1009 AATTGATCAATACACCAATATATGAGACTGCTTATCAAGCTCTTCTAGACATTTACAG 1068
QY 1037 AGAATGGAAGATTTCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1096
Db 1069 TGAATGGAACAGTGTGCTCAAGAGAGGTAACTGGACCTGTATATCTATGCAAAAAA 1128
QY 1097 ATTGTGAAGAGTGTGTTAGAAACCTGATGTTGAAGCAAAATGGCAAAATGA--GGG 1153
Db 1129 TGAGATGAAAAGTGTGTTGAGGCTTATTTAAGGAACCCATGTTGATGATGTA 1188
QY 1154 ACATATACCAACCACTGAAGAGAGATGATGAGTGTGATGATGATGATGATGATGATG 1213
Db 1189 CCATATTCAAAATATGAGGAACAAGTGGAGATGCAATCGTAAGTGTCTGCTATATGAT 1248
QY 1214 GCTTACAAACAACTGTGTTATCTTGGCATGATGATATATTCACAAAAGAGCTGTGGAATG 1273
Db 1249 GATATCAACAACTGTGTTGGTGGTATAGAGATTTATATCCACGAGACTTTTGAATG 1308
QY 1274 GGCTGTCTGTGACCTCTCTTTTATGATGATGATGATGATGATGATGATGATGATGATG 1333
Db 1309 GTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1368
QY 1334 TCTCATGACCCACAGCGGAGCAAGAAAGAAACATAGTTCATCGAGCTTTGAAAGTTA 1393
Db 1369 TATTGTTGACATGAAGATGAACAAAGAAAGAGACATGATGATGATGATGATGATGATG 1428
QY 1394 TATGAAGGAATATAATGTCAATGAGAGATATGCCAAACCTTGATTTACAGAGAGTAGA 1453
Db 1429 CATGAAGAGATTATGGAGCTTCAAGCAAGAGACTTACATTAAGTTCCTGAAAGAGTCA 1488

QY 1454 AGATGCTGTGGAAAGTATATAACCGAGAGTACCTCACAACATAAACAATTTCCAAAGCGTT 1513
Db 1489 CAATGCTATGGAAGGACATAAACAACAATTTCCCGTCCAACTGAAGTACCAATTTTGT 1548
QY 1514 ATTGATGCTGTGATCTATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1573
Db 1549 CTTTGAACAGTCTTAAATTTGACACGCTGGCTGACACGCTTATATATAAGGAGAGATAC 1608
QY 1574 CTTTCAACGATGAGGAGAGATACAAACATCTCATAAAGTCTCTACTCGTTTATCCTAT 1633
Db 1609 ATATTCAACCCCAAGGAAACCTTAAACAACATGATTAAATCCAACTACTAAATTGAT 1668
QY 1634 GAGTATATGAGGAT 1647
Db 1669 CAAAATATAAATAT 1682

RESULT 4
US-09-887-586A-51
; Sequence 51, Application US/09887586A
; Patent No. 6495354
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 64953541, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/887,586A
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 51
; LENGTH: 2024
; TYPE: DNA
; ORGANISM: Lycopodium esculentum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (32)...(1675)
; OTHER INFORMATION: VENT germacrene C synthase
US-09-887-586A-51

Query Match 21.7%; Score 357.6; DB 4; Length 2024;
Best Local Similarity 52.7%; Pred. No. 1.4e-95;
Matches 861; Conservative 0; Mismatches 749; Indels 24; Gaps 3;

QY 29 TCGCCCATTCGCAACTTCTCCCAAGCATTTGGGAGATCAGTTTCTCATCTATCAAAA 88
Db 58 TCGCCCATTCGCTAATTTTACCCTATGTTGGGATATCATTTCTTCTTATCTCA 117
QY 89 GCAAGTAGAGCAAGGGGTGGAAACAGATAGTGAATGATTAAAAAAGAGTGGCGCAACT 148
Db 118 TGAATTAATAATCAAGAAAAAGTTGAAGTTGATGATGATGATGATGATGATGATGATG 177
QY 149 ACTAAAGAAAGCTTTGGATATCTCTATGAACATGCCAATTTGTTGAAGCTGATTGATGA 208
Db 178 GCTGGTGAAGAACTGGGCAATAGCCTCAAAA-----GCTTGTGTTGATAGACGC 228
QY 209 AATTCAACGCTTGGATATACCGTATCACTTTGAACGGGAGATTTGATCATGCAATGCAATG 268
Db 229 GATGCAACGATTTGGAGTGGCTTATCTTTCGATAATGAAATTTGAACATCAATTTCAAAA 288
QY 269 TATTATGAACAT-----ATGGTGAATACTGGAAATGGTGAACCGCTCTCTCTT 316
Db 289 CATTTTGTGATCGTCCCAACAGAAATGATTAATGACAAACCTTTACGTTGTGTCTCT 348
QY 317 ATGGTTCGCTTATGCGAAGCAAGGATATTATGTTACATGATGATGATGATGATGATGATG 376

Db 349 TCCTTTTCGACTGTGAGGCAACAGGCCATTACATGCTTCAGATGTTCTCAAGCAATT 408
QY 377 TAAAGACAAAATAATGAGCGGTTCAAGCAATCGTTAGCTAATGATGTTGAAGGTTTCCTTGA 436
Db 409 CACCAACCAAGATGGGAAATTCAGAGAAACACTTACTTAATGATGTTCCAGGATTTATGAG 468
QY 437 GTTGTACGAACAACTTCTATGAGGGTACCTGGGGAGATTAATATAGAAAGATCTTTGG 496
Db 469 TTTGTATGAAGCATCATCTGAGAGTGGTAAATGAGGAGATTTCTTGAAGAAGCTCTTAC 528
QY 497 TTTTACACGATCTGCTCTAGCATTTATGACAAAGATGCTTTTCTACAAACCCGCTCT 556
Db 529 ATTACACCACTCATCTCGAGTCTATTGTCTCAACTTGGAGCAATAATAACTCTCT 588
QY 557 TTTTACCGAAATACACGGGCACTTAAAGCAACCCCTTTGGAAAAGGTTGCCAAGAAATAGA 616
Db 589 TAAGGTTGAAGTTGGTGAAGCTTAACTCAGCTTATTCGCACTTACCAAGATGGG 648
QY 617 GCGGGCGCATGATCTCTTCTATCAACAACAAGATCTCTATCAACAAGACTTTACTTAA 676
Db 649 AGCTAGAAAATACATATCCATTTACGAAACCAATGATGACACACCACTTTGCTTTGAA 708
QY 677 ACTTGCTAAGTTAGATTCATTTGCTTCACTGCTATTCGACAGGAGAGCTCAGCCATGT 736
Db 709 ATTGCTAAATGGATTTTAACTGCTGCAAAAGTTTACCAAAAGAGAGCTTAGTGATCT 768
QY 737 GTCAAAATGCTGAAAAGCTTTTCGATATCAAGAAGACGCACTTGTTTAAGAGATAGAA 796
Db 769 TACAAGGTGTGGAAGATTTGATTTTGGCAATAAATATCCATATGCAAGAGACAGTT 828
QY 797 TGTGTAATGCTACTTTTGGGAGCTAGGTTGAGGCTATGAGCCACAGTATTCGCGGCTAG 856
Db 829 GGTGAGTGTACTTCTCGATATTAGGAGTGTATTTTGAAGCAAAATATAGTCTGCGAG 888
QY 857 AGTTTCTTCAAAAAGCTTGTCTGTATTAACCTTATAGATGACACTTATGATGCTA 916
Db 889 AAAATGATGACAAAAGTACTCAACCTGACCTCCATTTATGACGACACTTTTGTGCTTA 948
QY 917 TGTACTTATGAAGAACTTAAAGTCTTACTGAGCTGTTGAAAGGTGGTCAATTACATG 976
Db 949 TGCAACCTTTGACGAATTTGCTGACTTTCAATGATGCAATCCAGAGATGGATGCTAATGC 1008
QY 977 CTTPAGACACTTCCAGAAATACATGAACCAACGATATACAAATATTTATGATGATACATAC 1036
Db 1009 AATTGATTCATCAACCACTATATGACACCTGCTTATCAAGCTCTCTTAGACATTTACAG 1068
QY 1037 AGAAATGGAAGATTTCTTCAAGAGGAGGAGAACAGATCTATTAACTCGCGCAAGA 1096
Db 1069 TGAATGGAACAAGTGTGTGCCAAGAAGGTAAACTGACCGTGTATATCTATGCAAAAAA 1128
QY 1097 ATTTGTGAAGAGTTTGTAGAAACCTGATGTTGAAGCAAAATGGGCAAAATGA---GGG 1153
Db 1129 TGNATGAAAAGTTGGTGAGAGCTTATTTAAGGAAACCCCAATGGTTGAATGTTGA 1188
QY 1154 ACACATACCAACCACTGAAGAGCATGATCCAGTTGTAATCATTTACTGGCGGTGTAACCT 1213
Db 1189 CCATATTCAAAATATAGAGAAACAAGTGGAGATGCAATCGTAAGTGTGCTATATGAT 1248
QY 1214 GCTTACAACTCTGTTATCTTGGCATGATGATATATTCACAAAAGAGTCTGCGAATG 1273
Db 1249 GATATCAACAACTTGTGTGCTGATAGAGATTTATATATCCACGAGACTTTTGAATG 1308
QY 1274 GGCTGCTCTGCACCTCTCTTTTATAGATCTAGGATATCTGCTGACGCCCTAAATGA 1333
Db 1309 GTTGATGAATGAGTCTGTGATTTTTCGAGCTTCCGCAATGANTCCAGAGCAATGAACGA 1368
QY 1334 TCTCATGCCCAAGCCGAGCAAGAAAGAAACATAGTTTCATCGAGCTTTGAAAGTTA 1393
Db 1369 TATTGTTGGACATGAAGATGAACAAGAAAGAGGACATGTAGCTTCACTTATTGATGTTA 1428
QY 1394 TATGAAGGAATATAATGTCAATGAGGAGTATGCCCCAACCTTGTATTTACAGGAGTGA 1453

Db 1429 CATGAAGATTTATGGAGCTTCAAGCAAGAGACTTACATTAAGTTCTGAAAAGAGGTAC 1488
QY 1454 AGATGTGTGGAAGATATATAACCGAGAGTACCTCACAACTAAAAACATTTCCAAGGCCGTT 1513
Db 1489 CAATGCATGGAAGGACATATAACAAACAATTTCTCCCGTCCAAGTGAAGTACCAATGTTGT 1548
QY 1514 ATTGATGGCTGTGATCTATTGTCGCAAGTTTCTTGAAGTTCAATATGCGAGGAAGGATAA 1573
Db 1549 CCTTGAACGAGTTCTAAATTTGACACGTTGCGCTGACACGCTTATATTAAGGAGAAATAC 1608
QY 1574 CTTCACACGTTATGGGAGACGAATACAAACATCTCAATAAGTCTCTACTCGTTTATCCTAT 1633
Db 1609 ATATTCAACCGCAAGGAACACTTAAAAACATGATTAATCCAATACTAATTGAATCTGT 1668
QY 1634 GAGTATATGAGGAT 1647
Db 1669 CAAAATATAAATAT 1682

RESULT 5
US-09-895-752-51
; Sequence 51, Application US/09895752
; Patent No. 6559297
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 65592971, Joseph P.
; APPLICANT: Starke, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/895,752
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 51
; LENGTH: 2024
; TYPE: DNA
; ORGANISM: Lycopersicon esculentum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (32)...(1675)
; OTHER INFORMATION: VENT germacrene C synthase
US-09-895-752-51

Query Match 21.7%; Score 357.6; DB 4; Length 2024;
Best Local Similarity 52.7%; Pred. No. 1.4e-95;
Matches 861; Conservative 0; Mismatches 749; Indels 24; Gaps 3;
QY 29 TCGCCCCATTCGCAACTTTCCCTCCAGAGATTTGGGAGATCAGTTTCTCATCTATCAAAA 88
Db 58 TCGCCCCCTTGGCTAATTTTCCACCCTCTGTTGGGATATCATTTCTTCTTATATCA 117
QY 89 GCAAGTGAAGCAAGGGTGGAAACAGATGATGAATGATTAATAAAAAAGAGTGGGCAACT 148
Db 118 TGAATTTACTAATCAAGAAAAAGTTGAAGTGTGATGATACAAAGAGACAATTAGAAAAAT 177
QY 149 ACTAAAAAGAGTTTGGATATTCCTATGAAACATGCCAATTTTGTGAGCTGATGATGA 208
Db 178 GCTGTGGAATTCGCAATAGCACTCAAAA-----GCTTGTGTTGATAGACGC 228
QY 209 AATTCAACGCTTGGAAATACCGTATCATTGCAACGGGAGATGATCATGCATTGCAATG 268
Db 229 GATGCAACGATTCGAGTGGCTTATTCATTTCCGATTAATGAAATGAAACATCCATTCAAA 288
QY 269 TATTATGAAACAT-----ATGGTGATAACTGGAATGGTGACCGCTCTTCTCTT 316

Db 289 CATTTCCTGATCGTCGTCACCAACAGATGATATGATGACCAACCTTTACGTTGTGCTCT 348
QY 317 ATGTTCCGTTTATGCGAAGCAGGATATTTATGTTACATGATGTTTTCATTAACHTA 376
Db 349 TCGTTTCGATTTGAGGCAACAGGCAATTAACATGCTTCAGATGTTTCAAGCAAT 408
QY 377 TAAAGACAAAATGAGCGTTCAGCAATCGTTAGCTAATGATGTTGAAGGTTTGCTTGA 436
Db 409 CACCAACCAAGATGGAATTCAGGAAACACTTACTAATGATGTTCCAGAGTATTGAG 468
QY 437 GTTGACCAAGCAATCTTATGAGGTTACCTGGGAGATTATATAGAAAGATGCTCTGG 496
Db 469 TTTGTATGAAGCATCACATCTGAGAGTGGTAATGAGGAGATTCITGAAGAGCTCTTAC 528
QY 497 TTTTACAGATCTGCTTAGCATTTATGACAAAAGATGCTTTTTCTACAAACCCGCTCT 556
Db 529 ATTTACCACTCTCTCGAGTCTATTGCTTCCAACTTTGAGCAATAATAAATCTCTCT 588
QY 557 TTTTACCAATAACAAACGGGCTAAAGCAACCCCTTTTGAAAAAGTTGCCAAGATAGA 616
Db 589 TAAGTTGAAGTTGGTGAAGCTTTAACTCAGCTATTCCATGACTTTACCAAGGATGG 648
QY 617 GCGGCGGAGTACATCTCTTCTATCAACAAAGATTTCTATACAAAGACTTTACTTAA 676
Db 649 AGCTAGAAAATACATATCCATTTACGAAAACAAATGATGACACACCACTTTGCTTTGAA 708
QY 677 ACTTGCTAAGTTAGATTTCAATTTCTGCTCAGTTCATTTGACAGGAGAGCTCAGCATCT 736
Db 709 ATTTGCTAAATGGATTTTAAATGCTGCAAAAGTTTCCAAAAGAGCTTAGTGATCT 768
QY 737 GTGCAAAATGGTGAAGCTTTTCTGATATCAAGAAAGACGACCTTTGTTTAAAGATAGAAAT 796
Db 769 TACAAAGTTGGTGAAGATTTGGAATTTGCAATAATAATCCATATGCAAGACAGGTT 828
QY 797 TGTGAATGCTACTTTTGGGAGTCTAGGTTACGGTATGAGCCACAGTATTCCTCGGCTAG 856
Db 829 GGTTCAGTGTACTCTCGATATTAGGAGTGTATTTGAGCCAAAATATAGTCGTCGAG 888
QY 857 AGTTTCTTCAAAAGCTTTGCTGTATTAATCTTATAGATGACACTTATGATCGTA 916
Db 889 AAAATGATGAAAAGTACTCAACCTGACCTCCATTTATGACGACACTTTTGATGCTTA 948
QY 917 TGGTACTTATGAAGAACTTAAGATCTTTACTGAAGCTTTGAAAGGTGTCATTAACATG 976
Db 949 TGCAACCTTTGACGAATCTGTGACTTTCAATGATGCAATCCAGAGATGGATGCTAATGC 1008
QY 977 CTTAGACACTTCCAGAAATCATGAACCGGATATACAAATTTATCATGATGATACAC 1036
Db 1009 AATTGATTCATACCAACCATATATGAGACCTGCTTATCAAGCTCTTCTAGACATTTACAG 1068
QY 1037 AGAATGGAAGATTTCTTGCAAGGAGGGAAGACAGATCTATTAACTGCGGCAAGA 1096
Db 1059 TGAATGGAACAGTGTGTCGAAAGAGGTAACTGGGACCGTGTATATCTATGCAAAAA 1128
QY 1097 ATTTGTGAAGAGTTGTTAGAAACCTGATGTTGAAGCAAAATGGGCAAAATGA---GGG 1153
Db 1129 TGAGATGAAAAGTTGGTGAGAGCTATTTTAAAGGAAACCAATGTTGATGATTGTGA 1188
QY 1154 ACACATCAACCACTGAAGACATGATCCAGTTGATATCATCTGCGGCTGCTAACCT 1213
Db 1189 CCATATCCAAAATATGAGGAAACAGTGGAGATGCAATCGTAAGTCTGGCTATATGAT 1248
QY 1214 GCTTACAACTCTGTTATCTTGGCATGATGATATTTCAAAAGAGTCTGTCCAAGT 1273
Db 1249 GATATCAACAACTTCTGTTGGTGGTATAGAAGATTTATCCAGGAGACTTTTGAATG 1308
QY 1274 GGCTGCTCTGCACTCTCTTTTATAGATCTCAGGTATATCTGTCGAGCGCTAAATGA 1333
Db 1309 GTTGATGATGATGTTGATGTTTCCGACTTCCGCAATTTGATCCAGAGCAATGACGA 1368
QY 1334 TCTCATGACCCCAAGGCGAGCAAGAAAGAAACATAGTTCTCATCGAGCCTTGAAGTTA 1393

Db 1369 TATTGTTGGACATGAAGATGAACAAGAAAGAGGACATGTAGCTTCACTATTGATGTTA 1428
QY 1394 TATGAGGATATAATGTCATGAGGAGTATGCCCAACCTTGTGATTTACAAGGAGTAGA 1453
Db 1429 CATGAAGATTATGAGCTTCAAGCAAGAGACTTACATTAAGTTCTCTGAAAGAGTCA 1488
QY 1454 AGATGTGTGAAAGATATAAAACCGAGAGTACCTCACAACTAAAAACATTTCCAAAGCCGTT 1513
Db 1489 CAATGATGAGGAGACATAAAACCAAACTCTCCGCTCCAACTGAAGTACCAATGTTGT 1548
QY 1514 ATTGATGGCTGTGATCTATTTTGTGCCAGTTTCTTGAAGTTCAANTATGCAAGGAGATAA 1573
Db 1549 CCTTGAACGAGTTCTAAATTTGACACGTTGGCTGACACGTTATATAAGGAGAAAGATAC 1608
QY 1574 CTTTACACGTTATGGGAGACGAATACAAACATCTCATAAAGTCTCTACTCGTTTATCTTAT 1633
Db 1609 ATATTCAACCGCAAGGAAAACCTTAAAAACATGATTAATCCAATACTAAATTGATCTGT 1668
QY 1634 GAGTATATGAGGAT 1647
Db 1669 CAAAATATAAATAT 1682

RESULT 6
US-09-903-012B-51
; Sequence 51, Application US/09903012B
; Patent No. 6569656
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 65696561, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/903,012B
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 51
; LENGTH: 2024
; TYPE: DNA
; ORGANISM: Lycopersicon esculentum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (32)...(1675)
; OTHER INFORMATION: VFNT germacrene C synthase
US-09-903-012B-51

Query Match 21.7%; Score 357.6; DB 4; Length 2024;
Best Local Similarity 52.7%; Pred. No. 1.4e-95;
Matches 861; Conservative 0; Mismatches 749; Indels 24; Gaps 3;

QY 29 TCGCCCAATGCCAATCTTCTCCCAAGCATTTGGGAGATCAGTTTCTCACTATCAAA 88
Db 58 TCGCCCAATGCCAATCTTCCCAATCTGTTGGGATATCATTTCTTCTTACTCA 117
QY 89 GCAAGTAGACCAAGGGGTGGAACAGATGATGAATTTAAAAAAGAGTGGCGCACT 148
Db 118 TGAATTAATTAACAAGAAAAGTTGAAGTTGATGAGTACAAAGAGACAAATAGAAAAT 177
QY 149 ACTAAAAGAGCTTTGGATATTCCTATGAACATGCCCAATTTGTTGAAGCTGATGAGA 208
Db 178 GCTGGTGAACCTTGCACAAATAGCACTCAAAA-----GCTTGTGTTGATAGAGC 228
QY 209 AATTCAACGCTTTGGATACCGTATCACTTTGAACGGGAGATTCATCATGATTCGAATG 268
Db 229 GATGCAACGATGGAGTGGCTTATCATTTTGCATTAATGAAATGAAACATCAATCAAAA 288

Qy 269 TATTATGAACAT-----ATGGTGATAACTGGAATGGTGACCGCTCTTCCTT 316
Db 289 CATTTTTCATGTCATCGTCCAAACAGAAATGATAATGACAAACACCTTTTACGTTGTGCTCT 348
Qy 317 ATGGTTCGCTTATGCGAAGCAGGATATTTATGTTACATGATGTTTTCNAATACTA 376
Db 349 TCGTTTTCGACTTGTGAGGCAACAAAGGCCATACATGCTTCCAGATGTTTTCARCAANT 408
Qy 377 TAAAGACAAAATGAGCGTTCAGCAATCGTTAGCTAATGATGTTGAAGGTTTGCTTGA 436
Db 409 CACCAACCAAGATGGGAATTCAGGAACACCTTACTATGATGTTCCAGAGATATTGAG 468
Qy 437 GTTGACGAACCACTTCTATGAGGATACCTGGGAGATTAATTATAGAGATGCTCTTGG 496
Db 469 TTTGTATGAAGCATCACATCTGAGAGTGGTAATGAGGAGATTCCTTGAAGAGCTCTTAC 528
Qy 497 TTTTACAGATCTGCTCTTAGCAATATGACAAAGATGCTTTTCTACAAAACCCGCTCT 556
Db 529 ATTTACCAACCACTCATCTCGAGTCTATTGTCTCCAACTTGGCAATAATAAATCTCT 588
Qy 557 TTTTACCAAAATACAAACGGGCACTAAAGCAACCCCTTTGGAAGGTTGCGCAAGATAGA 616
Db 589 TAAGTTGAAGTTGGTGAAGCCTTAACTCAGCCTATTCGATGACTTTTACCAAGGATGGG 648
Qy 617 GCGGCGCAGTACATTCCTTTCTATCAACAAGATTTCTCATACAAGACTTTTACTTAA 676
Db 649 AGCTAGAAAATACATATCATTTACGAAAACAATGATGATGACACCACTTTCGTTTGA 708
Qy 677 ACTTGCTAAGTTAGATTCATTTGCTTCAGTCAITGACAAAGAGAGCTCAGCCATGT 736
Db 709 ATTTGCTAATTTGGATTTTAAATGCTGCAAAAGTTTCACAAAGAGAGCTTAGTGATCT 768
Qy 737 GTGCAATGTTGGAAGCTTTTCGATATCAAGAAAGACGACCTTTTAAAGATAGAAT 796
Db 769 TACAAGTTGGTGAAGATTTGGATTTTCCAAATAAATATCATATGCAAGAGAGGTT 828
Qy 797 TGTGAAATGCTACTTTTGGGAGTACTAGTTTCAAGCTATGAGCCACAGTATTCGCGGCTAG 856
Db 829 GGTTCAGTGTACTCTGATATTTAGAGTGTATTTTGGAGCCAAATATAGTCTGCGGAG 888
Qy 857 AGTTTCTTCAAAAAGCTGTGTTTATTAATCTTATAGATGACACTTATGATCGGTA 916
Db 889 AAAAATGATGACAAAAGTACTCAACCTGACCTCCATTTATGAGGACACTTTTATGCTTA 948
Qy 917 TGGTACTTATGAAGAACTTAAAGATCTTTTCAAGCTGTTGAAAGTGGTCAATTAATG 976
Db 949 TGCAACCTTTGACGAACCTTGTGACTTTCAATGATGCAATCCAGAGATGGATCTAATGC 1008
Qy 977 CTTAGACACACTTCCAGAAATACATAAACCAGATATACAAATTTATTCATGATACATAC 1036
Db 1009 AATTGATTCAATACAAACCATATATGAGACCTGCTTATCAAGCTTTCTAGACATTTACAG 1068
Qy 1037 AGAAATGGAAGATTTCTTGCAAAAGAGGAGAGAACAGATCTATTTAACTGCGGCAAGA 1096
Db 1069 TGAAATGGAACAAAGTGTGTCCAAAGAGGTAACTGGGACCGGTATATCTATGCAAAAA 1128
Qy 1097 ATTTGTGGAAGATTTGTTAGAAACCTGTGTTGAAGCAAAATGGGCAATGA---GGG 1153
Db 1129 TGAGATGAAAAGTTGGTGAGAGCCATTTTAAAGGAACCCAAATGGTGAATGTTGTA 1188
Qy 1154 ACACATACCAACCACTGAAAGAGCATGATCCAGTTGTAATCATATPACTGGGGTCTAACCT 1213
Db 1189 CCATATTCAAAATATGAGGAACAAGTGGAGATGCAATCGTAAGTCTGGCTATATGAT 1248
Qy 1214 GCTTCAACAACCTGTTATCTTGGGATGATGATATATTCACAAAAGATCTCTCGAATG 1273
Db 1249 GATATCAACAACCTGCTTGGTGGTATAGAAATTTATATCCACAGAGACTTTTGAATG 1308
Qy 1274 GCCTGCTCTGCACTCTCTTTTATAGATACTCAGGTATATCTTGGTGGCGCCTAATGA 1333
Db 1309 GTTGAATGATGATGTTGATTTCCGCTTCCGATTTGATTTGCCAGAGCAATGAACGA 1368

Qy 1334 TCTCATGACCCCAAGCCGAGCAAGAAAGAAACATAGTTCTATCGAGCCTTGAAGTTA 1393
Db 1369 TATTGTTGGACATGAGATGAACAAGAAAGAGGACATGAGCTTCACTTATTGATGTTA 1428
Qy 1394 TATGAAGGAATATAATGTCAATGAGGAGTATGCCAAACCTTGATTTACAAGGAAGTAGA 1453
Db 1429 CATGAAGAATATGAGACTTCAAGCAAGAGACTTACATTAAAGTTCTCTGAAGAGGTAC 1488
Qy 1454 AGATGTGTGAAAGATATAAACCCGAGAGTACCTCACAACTAAAAACATTTCCAAAGGCCGTT 1513
Db 1489 CAATGCAATGAAGGACATATAACAAACAATTCCTCCGTCCAACTGAAGTACCAATGTTTGT 1548
Qy 1514 ATTGATGGCTGTGATCTATTTGTGCCAGTTTCTTGAAGTTCAATATGCAAGGAAGATAA 1573
Db 1549 CCTTGAACGAGTTCTATAAATTTGACACGTGTGGTGACACGTTATATAAGGAAAGATAC 1608
Qy 1574 CTTCAACAGTATGGGACAGCAATACAAACATCTCATAAAGTCTCTACTCTGTTTATCTCTAT 1633
Db 1609 ATATTCAACCGCCAAAGGAACCTTAAAAACATGTTTAACTCAATTAATTGAATCTGT 1668
Qy 1634 GAGTATATGAGGAT 1647
Db 1669 CAAAATATAAATAT 1682

RESULT 7
US-09-900-797-51
; Sequence 51, Application US/09900797
; Patent No. 6645762
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 66457621, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/900,797
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Fast-Seq for Windows Version 3.0
; SEQ ID NO 51
; LENGTH: 2024
; TYPE: DNA
; ORGANISM: Lycopersicon esculentum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (32)...(1675)
; OTHER INFORMATION: VFNT germacrene C synthase
US-09-900-797-51

Query Match 21.7%; Score 357.6; DB 4; Length 2024;
Best Local Similarity 52.7%; Pred. No. 1.4e-95;
Matches 861; Conservative 0; Mismatches 749; Indels 24; Gaps 3;
Qy 29 TCGCCCATTCGCCAACTTCTCCCAAGCATTTGGGAGATCAGTTTCTCATCTATCAAAA 88
Db 58 TCGCCCTTGGCTAATTTTCAACCATCTGTTTGGGATATCATTTCTTCTATACTCA 117
Qy 89 GCAGTAGAGCAAGGGTGGACAGATAGTGAATTTAAAAAAGAGTGGCGCAACT 148
Db 118 TGAATTAATACAGAAAAGTTGAAGTTGATGATACAAAGAGACAAATTAGAAAAT 177
Qy 149 ACTAAAAAGAGCTTTGGATATTTCCCTATGAAAACATGCCAAATTTGTTGAAGCTGATGA 208
Db 178 GCTGGTGAACCTTTCGACATAGCACTCAAAA-----GCTTGTGTTGATAGACGC 228
Qy 209 AATCAACGCCCTCGAATACCGTATCACTTTGAACGGGAGATGATCATGCATTGCAATG 268

Db 229 GATGCAACGATTGGGAGTGGCTTATCATTTTCGATATGAATTTGAAACATCCATTCAAAA 288
Qy 269 TATTATGAAACAT-----ATGGTGAATGAAATGGTGAAGCTTCTCTT 316
Db 289 CATTTTTCATGATCGTCCAAACAGAAATGATGACAAACCTTTTACGTTGTGTCCT 348
Qy 317 ATGTTTCGCTTATGCGAAACCAAGGATATTATGTTATCATGATGATGTTTCAATAACTA 376
Db 349 TCGTTTCGACTTGTGAGCAACACAGGCCATTCATGCTTCAGATGTTGTCAGCAAT 408
Qy 377 TAAAGACAAAATGAGCGGTTCAAGCAATCGTTAGCTAATGATGTTGAAGTTGCTTGA 436
Db 409 CACCAACCAAGATGGGAAATTCAGGAAACACTTACTAATGATGTCGAAGGATTTAGAG 468
Qy 437 GTTGTACAGCAACTTCTATCAGGGTACCTGGGAGATTTATAGAGATGCTCTTGG 496
Db 469 TTTGTATGAAGCATCACATCTGAGAGTGCCTAATGAGGAGATTCCTTGAAGAGCTCTTAC 528
Qy 497 TTTTACAGCATCTGCTCTTAGCATTTATGACAAAGATGCTTTTCTACAAACCCGCTCT 556
Db 529 ATTTACACCACTCATCTCGAGTCTATTGCTCCAACTTGAGCAATTAATAACTCTCT 588
Qy 557 TTTTACGAAATPACACGGGCACTAAAGCAACCCCTTTGGAAGAGTTGCCAAGATAGA 616
Db 589 TAAGGTTGAAGTTGGTGAAGCTTAACTCAGCCTATTGCGATGCTTTACCAAGGATGG 648
Qy 617 GCGCGCGAGTACATCTCTTCTATCAACCAAGATTTCTCATCAAGACTTTACTTAA 676
Db 649 AGCTAGAAATACATATCATTTACGAAACCAATGATGCACACCACTTTGCTTTGAA 708
Qy 677 ACTTCTTAAGTTAGAGTTCAATTTGCTTCAGTCAATTCACAGGAAGAGCTCAGCCATGT 736
Db 709 ATTTCTTAATTTGGATTTTAAATGCTGCAAGTTTCCAAAGAGAGCTTAGTATCT 768
Qy 737 GTGCAATGCTGAAAGCTTTGATATCAAGAAAGAGCAAGCTTTTGAAGATAGAT 786
Db 769 TACAAGGTTGGAAGATTTGGATTTGCAATAAATATCCATATGCAAGAGACAGGTT 828
Qy 797 TGTGATGCTACTTTTGGGACTAGTTTCAGCTATGAGCCACAGATTTCCCGGCTAG 856
Db 829 GGTGAGTGTACTTCTGATATAGAGTGTATTTGAGCCAAATATATAGTCTGTCGAG 888
Qy 857 AGTTTCTTCAAAAAGCTTGTCTGTATTAATCTTATAGATGACACTTTATGATGCTA 916
Db 889 AAAATGATGACAAAGTACTCACTGACCTCCATTTATGACGACACTTTGATGCTTA 948
Qy 917 TGGTACTTATGAAGAACTTTAAGATCTTTTACTGAGCTTTGAAAGTGGTCAATTAATG 976
Db 949 TGCACCTTTGACGAACTTGTGACTTTCAATGATGCAATCCAGAGATGGATGCTAATGC 1008
Qy 977 CTTAGACACACTTCCAGATACATGAAACCGATATACAAATTTATCATGATACATACAC 1036
Db 1009 AATTGATTCATACAACTATATGAGCTGCTTTATCAAGCTCTTCTAGACATTTACAG 1068
Qy 1037 AGAAATGGAAGAAATTTCTTCAAGAGGAGGAACAGATCTTTTAACTGCGCGCAAGA 1096
Db 1069 TGAATGGAACAGTGTGTCGAAGAGGTAACCTGGACCGTGTATCTATGCAAAAA 1128
Qy 1097 ATTTGGAAGAGTTTGTAGAAACCTGATGTTGGAAGCAAAATGGGCAATGA---GGG 1153
Db 1129 TCAGATGAAAAAGTTGGTGAGAGCTTATTTAAGGAAACCAATGGTTGAATGATGTA 1188
Qy 1154 ACACATACCAACCACTGAGAGCATCATCAGTTGTTGATCATCTACTGGCGGTGCTAACCT 1213
Db 1189 CCATATTCAAAATATGAGAAACAAGTGGAGATGCAATGTAAGTCTGGCTATATGAT 1248
Qy 1214 GCTTACCAACAACTGTTTCTTGGCATGATATATTCACAAAAGAGTCTGTGCAATG 1273
Db 1249 GATATCAACAACTTGTGTCGGTATAGAGAAATTTATATCCACAGACTTTTGAATG 1308
Qy 1274 GCGTGTCTCTGCAACCTCTCTTTTATAGATACTCAGGTATATCTGGTCGAGCCCTAAATGA 1333

Db 1309 GTTGATGAATGAGTCTGTGATTTGTTGAGCTTCGCGATTCGCGAGCAATGAACGA 1368
Qy 1334 TCTCATGACCCACAGGCCGAGCAAGAAAGAAACATAGTTTCATCGAGCCTTTGAAGTTA 1393
Db 1369 TATTGTTGGACATGAAGATGAACAAAGAAAGAGACATGACTTCTTCACTTATGAATGTTA 1428
Qy 1394 TATCAAGGAATATAATGTCAATGAGGAGTATGCCCAACCTTTGATTTTCAAGGAAGTAGA 1453
Db 1429 CATGAAGATTTAGGAGCTTCAAGCAAGAGACTTACATTAAGTTTCTCTGAAAGAGTCAAC 1488
Qy 1454 AGATGTTGGAAGATATAAACCGAGAGTACTCTCAAACTAAAAACATTTCCNAGGCCGTT 1513
Db 1489 CAATGATGGAAGGACATAAACAACAATTTCTCCGCTCCAACTGAAGTACCAATGTTGT 1548
Qy 1514 ATTGATGCTGTGATCTATTGTTGTCAGTTTCTTTGAAGTTCAATATGAGGAAGGATAA 1573
Db 1549 CTTTGAAGAGTCTTAATTTTGACACGCTGGCTGACAGCTTATATAAGGAAGAGATAC 1608
Qy 1574 CTTTCAACGATGAGGAGCAATAACAACATCTCATAAAGTCTCTACTCGTTTATCTCTAT 1633
Db 1609 ATATTCAACGCCCAAGGAAACCTTAAAAACATGATTAAATCAATTAATTTGAATCTGT 1668
Qy 1634 GAGTATATGAGGAT 1647
Db 1669 CAAAATATAAATAT 1682

RESULT 8

US-09-398-395A-31
; Sequence 31, Application US/09398395A
; Patent No. 6468772
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 6468772, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/398,395A
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 1944
; TYPE: DNA
; ORGANISM: Solanum tuberosum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (57)...(1724)
; OTHER INFORMATION: vetispiradiene synthase
US-09-398-395A-31

Query Match 20.0%; Score 330.4; DB 4; Length 1944;
Best Local Similarity 52.0%; Pred. No. 1.5e-87; Indels 30; Gaps 4;
Matches 857; Conservative 0; Mismatches 761;

Qy 14 AGAAGAAAAACCTATTCGCCCCCATTTGCCAACTTTCTCCAAAGCATTTGGGGAGATCAGTT 73
Db 92 AGAGAGGAGATTGTTTCGCCCCCATGCTGACTTCTCTCCAACTCTTTGGGGTATGTTT 151
Qy 74 TCTCATCTATCAAAAGCAAGTAGAGCAAGGGGTGAAACAGATAGTAATGATTTAAAAAA 133
Db 152 TCATTCTTCTCCCTCGACAATCAGATTGCTGAAAAATATGCTCAAGAGATCGAAACTTT 211
Qy 134 AGAAGTGGGCACTACTATAAAGAGCTTTGGATATTCTATGAACATGCCAATTTGTT 193
Db 212 GAAGGAACAATCAAGAATTATATTATCTGCACTCTTCTCGAAGAACATTCGCTGAGAAAT 271

194 GAAGCTGATTGATGAATTCAGCGCTTGGAAATACCGTATACCTTTGAAACGGGAGATTGA 253
Db
272 GGATCTGATAGACATGTTGAGCGCTTGGCGATGCTTATCATTTTGAAAAACAATAGA 331
Qy
254 TCATGCAATGCAATGATTATTAAGAACATATGCT-----GATAACTGGAA 298
Db
332 TGATATGTTGGATCAATTTTCAAGACAGATCTTAATTTGAGGCTCACAGATCAATGA 391
Qy
299 TGGTGACCGCTCTTCTTATGTTCCGTTCCGTTTATGCGAAAGCAAGATATTTATGTTACATG 358
Db
392 TTTACAAACTTTATCCGTTCAATTTTCGACTATTGAGACAACATGTTTACAATATCTCCCC 451
Qy
359 TGATGTTTTCAATACTATAAGACAAATATGGAGGTTCAAGCAATCTGTAGCTAATGA 418
Db
452 AAAACTTTTTATFAGATTCCAAGATGCAAAAGGCAAAATTTAAGAAATCTCTTTGTTAACA 511
Qy
419 TGTGAAGGTTTGCTTTGTTGTTAGTTGTAAGCAACTTCTATGAGGTTACCTGGGAGATTAT 478
Db
512 CATCAAGGCTCTTTGAATTTATGAAGCTTCGATGTGAAGCTCATGGAGATAT 571
Qy
479 ATTAGAAGATGCTTTGTTTTTACAGATCTCGTCTTAGCATTTATGACAAAAGATGCTTT 538
Db
572 TTTGGAAGGCACTTGTCTTCTACTGCTCATCTTGAATCT-----GCAGCTCC 622
Qy
539 TTCTACAAACCCGCTCTTTTACCGAAATACACGGGCACTAAAGCAACCCCTTTGGAA 598
Db
623 ACATTTGAAGTCACTCTGAGTAAGCAAGTGCACATGCGCTTGAGCAATCTCTCCATAA 682
Qy
599 AAGGTTGCCAAGATFAGAGGGCGGCGAGTACATCTCTTCTATCAACACAGATTTCTCA 658
Db
683 GAGCATTCAGAGTTGAGACAGGCTACTTCTATCTATCTATACGAGAGGAGAAACAGAA 742
Qy
659 TAACAAGACTTTACTAAACTTGTCTAAGTTAGAGTTCAATTTGCTTCAGTCAATGACAAA 718
Db
743 GAATGATGTTGCTTCAATTTGCAAACTGCACTTCACTTACTTACAGATGTGACAA 802
Qy
719 GGAAGAGCTCAGCCATGTTGCAAAATGTTGGAAGCTTTCGATATCAAGAAAGCCACC 778
Db
803 ACAAGAATCTTAGTGAAGTATCAAGGTGTTGGAAGATTTGGATTTGTGACACACTTCC 862
Qy
779 TTGTTTAGAGATAGAAATTTGTTGAATGCTACTTTTGGGAGCTAGGTTCAAGGCTAGAGC 838
Db
863 ATATGTTAGGATAGAGCAGTGGAAATGCTACTTTTGGAGATGGGGGTGTATGCTGAAC 922
Qy
839 ACAGTATTCGGGCTAGAGTTTCTTCAAAAGCTGTTGCTGTTTAACTCTTATAGA 898
Db
923 TCAATACTCTCAGGCTCGTGTCATCTGCTAAGACTATAGCAATGATTTCTATAGTAGA 982
Qy
899 TGACACATTTGATGGTATGTTATGAGAACTTTAAGAACTTTTACTGAAAGCTGTTGA 958
Db
983 TGACACATTCGATGCTTATGTCATTTGTCAAAAGAACTTGAGATCTACACCGATGCCATA 1042
Qy
959 AAGTGGTCAATTTACATGCTTAGACACACTTCCAGNATACATGAACCGATATACAAAT 1018
Db
1043 GAGTGGGATATTAGCCAAATTTGATCGGCTCCCTGATTACATGAATACTAGTTACAAAG 1102
Qy
1019 ATTATGATATACATACACAGAAATGGAAGAAATTTCTTCAAAAGGAGGGAAGAACAGATCT 1078
Db
1103 ACTTTAGATCTCTACATGATTATGAATGAGTTGTCCAAGGATGGTAGATCTGATCT 1162
Qy
1079 ATTAACTGCGGCAAGAAATTTGTGAAGAGTTGTTGAAGAACTGATGTTGGAAGCAAA 1138
Db
1163 TGTTCACCTACCGCAAGAAAGAAATGGAAGAAATCTGTGAAACTATTTTGTGGAAGCAA 1222
Qy
1139 ATGGGCAATGAGGACACATACCAACCACTGAAGAGCATGATCCAGTTGATATCATTTAC 1198
Db
1223 ATGTTTCAATGAAGATATATGCCGCCAGTCTCTGAGTATCTTAGCAATGCAATGATGCTAC 1282
Qy
1199 TGGCGGTCTAACTGCTTACAACTGTTTATCTTGGCATGAGTGATATATTCACAAA 1258
Db
1283 CAGCACTTATTACTTGTCTACACTACATCTTATTTGGCATGAAGTCTGCTAACAGCA 1342

Qy 1259 AGAGTCTGTGCAATGGGCTGTCTCTGCACCTCTCTTTTATAGATACTCAGGTATCTTGG 1318
Db 1343 AGATTTTG---AATGGTTGGCCAAAGAACCTTAAATTTCTTGAGGCTTAATGTGACGTTATG 1399
Qy 1319 TCGAGCGCTTAAATGATCTCATGACCCCAAGCCGAGCAAGAAAGAAACATAGTTCATC 1378
Db 1400 CCAGTCTATAGATGACATAGCCCTATGAGGTTGAGAAGGGTAGAGTCAATTTGCCAC 1459
Qy 1379 GAGCCTTTGAAAGTTATATGAAGGAATATAATGTCTCAATGAGGAGTATGCCCAACCTTGAT 1438
Db 1460 TGAATTTGAATGTTTACATGAGAGATTATGGTGTATCCACAGAAAAGGCCATGAAAAAT 1519
Qy 1439 TTACAGGAAGTAGAAGATGTTGGAAGATATAAAACGAGAGTACCTCACAACTTAAAAA 1498
Db 1520 CCAAGAAATGGCTGAGACAGCATGGAAGGATGTAATGAAGGAATCTTCGACCAACTCC 1579
Qy 1499 CATTTCCAAAGCCGTTATTGATGGCTGTGATCTATTGTCAGATTTCTTGAAGTTCAATA 1558
Db 1580 CGTCTCTACAGAGATCTCTACTGCAATCTCAATCTTGCTCGCATTTATCGATGTTACTTA 1639
Qy 1559 TGCAGGAA---AGGATAACTTTCACAGTATGGGAGACGAATACAAACATCTCATAAAGTC 1615
Db 1640 TAAGCACAATCAAGATGGATACACTCATCGGAAAAAGTACTTAAACCTCATATATTATGC 1699
Qy 1616 TCTACTCGTTTATCTCTATGATATATGA 1643
Db 1700 GTTGTGGTGGACTCTATTGAAATTTAA 1727

RESULT 9

US-09-887-586A-31
; Sequence 31, Application US/09887586A
; Patent No. 6495354
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: NO.64953541, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/887,586A
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 1944
; TYPE: DNA
; ORGANISM: Solanum tuberosum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (57)...(1724)
; OTHER INFORMATION: vetispiradiene synthase
US-09-887-586A-31

Query Match 20.0%; Score 330.4; DB 4; Length 1944;

Best Local Similarity 52.0%; Pred. No. 1.5e-87;

Matches 857; Conservative 0; Mismatches 761; Indels 30; Gaps 4;

Qy 14 AGRAGAAAACCTTATTCGCCCATTCGCCACTTTCTCCRAGCATTTTGGGAGATCAGTT 73
Db 92 AGAGGAGGAGATTGTTGCCCATAGCTGCTTCTCCAAGTCTTTGGGGTATCGTTT 151
Qy 74 TCTCATCTATCAAAAGCAAGTAGAGCAAGGGGTGGAACAGATAGTGAATGATTTAAAAA 133
Db 152 TCATTATCTCCCTCGACATCAGATTGCTGGAATATATGCTCAAGAGATCGAACTTT 211
Qy 134 AGAAGTGGCGCAACTACTATAAAGAGCTTTGGATATTCTCTATGAACATGCAATTTGTT 193

Db 212 GAAGGAACAATCAAGAAATTAATATCTGATCTTCTGGAAGAACATTTGGCTGGAAT 271
QY 194 GAAGCTGATGATGAATTCACGCGCTTGAATACCGTATCACCTTTTGAACGGGAGATTGA 253
Db 272 GGATCTGATAGACATTTGAGCGCTTGGCAITTCATTTTGAACCAATAGA 331
QY 254 TCATGCAATTCATGATTTATGAACATATGGT-----GATACTGGAA 298
Db 332 TGATATGTTGGATCAATTTTCAAGCAGATCCCTAACTTTGAGGCTCACGAGTACAATGA 391
QY 299 TGGTGACCGCTCTCTCTATGTTTCCGTCTTATGCGAAAGCAAGATATATGTTATGATG 358
Db 392 TTTACAACTTTATCCGTTCAATTTGACATTTGAGACACATGTTTACAAATATCTCCCC 451
QY 359 TGATGTTTCAATPAACATATAAGACAAAATGAGCGCTTCAAGCAATCGTTAGCTAATGA 418
Db 452 AAAACCTTTTATAGATTCACAGATGCAAAAGGCAAAATTTAAAGAACTCTCTTTGTAACGA 511
QY 419 TGTGAGGCTTGTCTGAGTTGACGAGTACGAGCAACTTCTATGAGGTACTCTGGGAGATAT 478
Db 512 CATCAAGGCTCTTTGAACTTATACGAAAGCTCGCATGTAAGGACTCATGGAGAGATAT 571
QY 479 ATTAGAAGATGCTCTCTGTTTACAGATCTCGTCTTAGCATATGACAAAAGATGCTTT 538
Db 572 TTTGAGAGGCACTTGTCTTCTACTGCTCACTTGAATCT-----GCAGCTCC 622
QY 539 TTCTACAAACCCCGCTCTTTTTCACGAAATPACACGGGCACTAAAGCAACCCCTTTGGAA 598
Db 623 ACATTTGAAGTCACTCTGAGTAAGCAAGTGACACATGCGCTTGAGCAATCTCTCCATAA 682
QY 599 AAGGTTGCCAAGATAGAGCGCGGAGTACATCTCTTCTATCAACAAAGATTTCTCA 658
Db 683 GAGCATTCCAAGAGTTGAGACAGCTACTTCACTCTATCTATCTAGGAAGGAGGAACAGAA 742
QY 659 TAACAGACTTTACTTAACTCTGTAAGTTAGAGTTCAATTTGCTCAGTCATTTGCACAA 718
Db 743 GAATGATGTTGCTTCAATTTCCAAACTGAGACTTCACTTCACTTCACTGATTTGACAA 802
QY 719 GGAAGAGCTCAGCCATGTGTGCAAAATGGTGGAAGCTTTTCATATCAAGAAAGCGCACC 778
Db 803 ACAAGAAGCTTAGTGAAGTATCAAGGTGGTGGAAGATTTGGAATTTGTGACACACITCC 862
QY 779 TTGTTTAAGAGATAGAAATTTGAATGCTACTTTTGGGAGCTAGTTTCAGGCTATGAGCC 838
Db 863 ATATGCTAGGGATAGACAGTGGAAATGCTACTTTTGGACGATGGGGGTGATGCTGAACC 922
QY 839 ACAGTATTCGGGCTAGAGTTTCTTTCACAAAGCTTTGCTGTTTAACTCTTATAGA 898
Db 923 TCAATACCTCAGGCTCGTGTGATGCTTGAAGACTATAGCAATGATTTCTATAGTAGA 982
QY 899 TCACACTTATGATGCGTATGGTACTTATGAAGAACTTAAGATCTTTACTGAAGCTGTGA 958
Db 983 TGACACATTCGATGCTTATGGCATTGTCAAGAACTTGAGATCTACACCGATGCCATACA 1042
QY 959 AAGGTGGTCAATATACATGCTTAGACACATCTCAGAAATACATGAACCGGATATACAAAT 1018
Db 1043 GAGGTGGGATATTAGCCAAATTTGATCGGCTCCCTGATATACATGAATATCAGTATCAAGC 1102
QY 1019 ATTCTAGGATACATACACAGAAATGAAGAATTTCTTGCAAGGAGGGAAGACAGATCT 1078
Db 1103 ACTTTAGATCTCAATGATTAATGAATGAGGTTGTCCAGAGTGGTAGATCTGATG 1162
QY 1079 ATTAACTCGGCAAGAATTTGTGAAGAGTTTGTGAAGAACTGATGTTTGAAGCAAA 1138
Db 1163 TGTCTACTACGGGAAGAAGAAATGAAGAAATCGTGAGAAATCTATTTTGTGAAGCAAA 1222
QY 1139 ATGGGCAATGAGGGACACATACCAACCTAGAGCATGATCCAGTTGTAATCATAC 1198
Db 1223 ATGGTTTATTGAAGGATATATGCCCGGAGCTCTGAGTATCTTAGCAATGATAGCTAC 1282
QY 1199 TGGCGGTGCTAACCTCTTACAACTGTTTATCTTGGCATGAGTGATATATTTACAAA 1258

Db 1283 CAGCACTTATTACTTGTCTTACGACTACATCTTATTTGGGATGAAGTCTGCTAACAGCA 1342
QY 1259 AGAGTCTCTGAAATGGGCTGTCTCTGCACTCTCTCTTTTAGATACTCTCAGGTATCTTGG 1318
Db 1343 AGATTTTG--AATGTTGGCCCAAGAACCTTAAATTTCTTGAGGCTAAJGTGACGTTATG 1399
QY 1319 TCGAGCCCTAAATGATCTCATGACCCCAAGCCCGAGCAAGAAAGAAACATAGTTTCATC 1378
Db 1400 CCGAGTCTATAGATGATAGCCACCTATGAGGTTGAGAAAGGCTAGAGGTGAGATTGCCAC 1459
QY 1379 GAGCCTTCAAAAGTTATATGAAGGAATATAATCTCAATGAGGAGTATGCCCAACCTTGAT 1438
Db 1460 TGAATTTGAATTTTACATGAGAGATTTGTTATCCACAGAAAGCCATGGAATAAT 1519
QY 1439 TTACAAGGAAGTAGAAGATGTGTGAAAGATATAAACCGAGAGTACTCTCAACCTAAAAA 1498
Db 1520 CCAAGAAATGGCTGAGACAGCATGGAAGATGTAAATGAAGAAATCTCTCGACCACTCC 1579
QY 1499 CATTCGAAGGCGTTATTTGATGGCTGTGATCTATTGTTGCCAGTTTCTTGAAGTTCAATA 1558
Db 1580 CGTCTCTACAGAGATTTCTCACTCGCAATTTCAATCTTGTCTGCAATTTATCGATTTACTTA 1639
QY 1559 TGCAGGAA--AGGATAACTTTCACACGATGATGGAGACGAATACAAACATCTCATAAAGTC 1615
Db 1640 TAAGCACAATCAAGATGATACACTCATCCGGAAGAAAGTACTAAACCTCATATTTATTCG 1699
QY 1616 TCTACTCGTTTATCCTATGAGTATATGA 1643
Db 1700 GTTGTGGTGGACTCTATTGAAATTTAA 1727

RESULT 10

US-09-895-752-31
; Sequence 31, Application US/09895752
; Patent No. 6559297
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 65592971, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/895,752
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 1944
; TYPE: DNA
; ORGANISM: Solanum tuberosum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (57)...(1724)
; OTHER INFORMATION: vetispiradiene synthase
US-09-895-752-31

Query Match 20.0%; Score 330.4; DB 4; Length 1944;
Best Local Similarity 52.0%; Pred. No. 1.5e-87;
Matches 857; Conservative 0; Mismatches 761; Indels 30; Gaps 4;
QY 14 AGAAGAAAAACCTATTCGCCCATTTGCCAATTTCTCTCCAGCATTTGGGAGATCAGTT 73
Db 92 AGAGGAGAGATTTGTCGCCCATAGCTGACTTCTCTCCAGTCTTTGGGTGATCGTTT 151
QY 74 TCTCATCTATCAAGCAAGTAGAGCAAGGGGTGGAACAGATAGTGAATTTTAAAAA 133

Db 152 TCATTCTCCCTCGCAATCAGATTGCTGGAATAATGCTCAAGAGATCGAACTTT 211
QY 134 AGAGTGGGCAACTACTATAAAGAGCTTTGGATATTCCTAAGAACATGCGCAATTTGTT 193
Db 212 GAAGGAACAAATCAAGAAATATATATCTGCAATCTTCTCGAAGAAATTTGGCTGAGAAAT 271
QY 194 GAAGCTGATTGATGAATTTCAAGGCTTTGGAATACCGTATCACTTTGAAACGGGAGATTGA 253
Db 272 GGATCTGATAGACATTTGAGGCGCTTGGCATTCCTTATCATTTTGAACAAATAGA 331
QY 254 TCATGCAATTGCAATGATTTTATGAACATATGCT-----GATACTGGAA 298
Db 332 TGATATGTTGGATCAATTTTCAAGCAGATCTTAACCTTTGAGGCTCAGGATCAATGA 391
QY 299 TGGTACCGCTCTTCTTATGGTTCGCTCTTATGCGAAAGCAGAGATATTTATGTTACATG 358
Db 392 TTTACAAACTTTATCCGTTCAATTTGCACTATTTGAGCAACATGTTTCAATATCTCCCC 451
QY 359 TGATGTTTTCAATAACTATAAAGCAAAATTTGAGCGTTTCAAGCAATCGTTAGCTAAATGA 418
Db 452 ABAATTTTTATTAGATTCGAAGTGCAGAAAGGCAATTTAAGAAATCTCTTTGTAAGA 511
QY 419 TGTGAAGGTTGCTTGATTTGTACGAAGCAACTTCTATGAGGGTACCTGGGAGATTAT 478
Db 512 CATCAAGGCTCTTTGAACTTATACGAAGCTCGCATGTAGGACTCATGAGAGAGATAT 571
QY 479 ATTAGAAGATGCTCTTGGTTTTACAGATCTCGCTTTAGCATATGACAAAGATGCTTT 538
Db 572 TTTGGAAGAGGCACTTGCTTTCTCTACTGCTCATCTTGAATCT-----GCAGCTCC 622
QY 539 TTCTFACAAACCCCGCTCTTTTACGAATAACACGGGCACCTAAAGCAACCCCTTTGGAA 598
Db 623 ACATTTGAATGACCTCTGATGATGCAAGTGCACATGCCCCCTTGAGCAATCTCTCCATAA 682
QY 599 AAGGTTGCCAAGATAGAGGGGGGCGAGTACATTCCTTTCTATCAACCAAGATTCTCA 658
Db 683 GAGCATTTCCAGAGTGTAGACACGCTACTTCTCTCTATCTACGAGAGGAGGACAGAA 742
QY 659 TAACAAGACTTTACTTAACTGCTTAAGTTAGATTTCAATTTGCTTCACTAGTCAATTCACAA 718
Db 743 GAATGATGTGTTGTTCAATTTGCAAACTGGACTTCAACTTACTTCAAGATGTTTCCACAA 802
QY 719 GGAAGAGCTCAGCATGTGTCBAATGTTGGAAGCTTTTCGATATCAAGAAAGACGACC 778
Db 803 ACAAGAACTTAGTGAAGTATCAAGGTGTGGAAGATTTGGATTTTGTGCAACACTTCC 862
QY 779 TTGTTTAAGATAGAAATTTGTTGAATGCTATTTTGGGCACTAGGTTTCAAGCTATGAGCC 838
Db 863 ATATGCTAGGATAGAGCAGTGGATGCTACTTTTGGACGATGGGGGTGATGCTGAACC 922
QY 839 ACAGTATCCCGGCTAGAGTTTCTTCAAAAAGCTGTTGCTGTTTAACTCTTTATAGA 898
Db 923 TCAATACTCTCAGGCTCGTGTGATGCTTGTCTTAAGACTATAGCAATGATTTCTATAGTAGA 982
QY 899 TGACACTTATGATCGGTATGCTTATGAGAACTTAAAGTCTTTACTGAAGCTGTTGA 958
Db 983 TGACACTTATGCTTATGCTATGCTATGCTTCAAGAACTTGAAGTCTACACCGATGCCATACA 1042
QY 959 AAGGTGCTCAATTTACATGCTTTAGACACACTTCCAGATACATGAACCGATATACAAAT 1018
Db 1043 GAGGTGGATATTAGCCAAATTTGATCGCTCCCTGATTTACATGAATAATCAGTTACAAAGC 1102
QY 1019 ATTCAATGATATACACAGAAATTTGAAGAAATTTCTTCAAGGAGGGAAGCAAGATCT 1078
Db 1103 ACTTTTATGATCTCTACAAATGATTTAATGAAATGGAGTTGTCCAAGGATGTTAGATCTGATG 1162
QY 1079 ATTAACTGGGCAAGAAATTTGTAAGAGTTTGTGTAAGAACTGATGTTGAGCAAA 1138
Db 1163 TGTTCATCTACCGGAAGAAAGAAATGTAAGAAATCTGTGAGAACTATTTTGTGGAAGCAAA 1222
QY 1139 ATGGGCAATGAGGGACACATACCAACCACTGAAAGCATGATCCAGTTGTAATCATTTAC 1198

Db 1223 ATGTTTCATTGAAGGATATATGCGGCAGTCTCTGAGTATCTTAGCAATGCAATTAGTAC 1282
QY 1199 TGGCGGTGCTAACTGCTTACAACTTGTATCTTGGCATGAGTATATATTTACAAA 1258
Db 1283 CAGCACTTATTTACTTGGTCTAGCACTACATCTTATTTGGGCATGAAGTCTGCTTAAACAGCA 1342
QY 1259 AGACTCTGTCGAATGGGCTGCTCTGTCACCTCTCTTTTATAGTACTCAGGTATACTTGG 1318
Db 1343 AGATTTTG---AATGTTGGCCAAAGAACCTTAAATTTCTTGAGGCTAATGTGACGTTATG 1399
QY 1319 TCGACCCCTAAATGATCTCATGCCCCCAAGGCCGAGCAAGAAAGAAACATAGTTTCATC 1378
Db 1400 CCGAGTCATAGATGACATAGCCACTATGAGTTGAGAGGGTAGAGGTGAGATTGCCAC 1459
QY 1379 GAGCCTTGAAGTTATATGAAGAAATATATGATCAATGAGAGATATGCCAAACCTTGAT 1438
Db 1460 TGGAAATTTGAATGTTATACATGAGATTTATGTTGATCCACAGAAAGGCCATGGAAAAAT 1519
QY 1439 TTACAAAGGAATGAGAATGTTGGAAGATATAAACCCGAGAGTACCTTCAACACTAAAAA 1498
Db 1520 CCAGAAATGGCTGACACAGCATGGAAGGATGTAAATGAAGGAATCCTTCGACCACTCC 1579
QY 1499 CATTCGAAGCCGTTATTTATGCTGTGATCTATTTTGGCCAGTTCTTTGAAGTTCAATA 1558
Db 1580 CGTCTCTACAGATTTCTCACTCGCATTTCTCAATCTTGTCTGCGATTATCGATGTTACTTA 1639
QY 1559 TGCAGGAA---AGGATAACTTCAACGATGGAAGAGCAATACAAACATCTCATAAAGTC 1615
Db 1640 TAAGCAAAATCAAGATGGATACACTCATCCGAAAAAGTACTTAAACCTCATATTATTGC 1699
QY 1616 TCTACTCGTTTATCTCTATGATATATGA 1643
Db 1700 GTTGTGGTGGACTCTATTGAAATTTAA 1727

RESULT 11

US-09-903-012B-31
; Sequence 31, Application US/09903012B
; Patent No. 6589656
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 65896561, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/903,012B
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 1944
; TYPE: DNA
; ORGANISM: Solanum tuberosum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (57)...(1724)
; OTHER INFORMATION: vetispiradiene synthase
US-09-903-012B-31

Query Match 20.0%; Score 330.4; DB 4; Length 1944;

Best Local Similarity 52.0%; Pred. No. 1.5e-87;

Matches 857; Conservative 0; Mismatches 761; Indels 30; Gaps 4;

QY 14 AGAAGAAAAACCTATTTCGCCCATTCGCAACTTTCTCCCAAGCATTTGGGAGATCAGTT 73

Db 92 AGAGGAGGAGATGTTTCGCCCATAGCTGACTTCTCTCCAAAGTCTTTGGGATCGTTT 151

QY 74 TCTCATCTATCAAAAGCAAGTAGCAGCAAGGGTGGAAACAGATAGTGAATGATTTAAAAA 133
Db 152 TCATTCATCTCTCCCTCGCAATCAGATGCTGGAAATATGCTCAAGAGATCGAACTTT 211
QY 134 AGAAGTCGGGCACTACTAAAAAGAGCTTTGGATATTCCTATGAAACATGCCAATTTGTT 193
Db 212 GAAGGAACAATCAAGAAATATATATCTGATCTTCTCGAAGAACATTTGCTGAGAAAT 271
QY 194 GAAGCTGATGATGAATTAACAGCCCTTGGAAATACCGTATACCTTTGAACGGGAGATTTGA 253
Db 272 GGATCTGATAGACATTTGAGGCGCTTGGCAATGCTTATCATTTTGAAGAAACAAATAGA 331
QY 254 TCATGCAATTCGATGATTTATGAAACATATGTT-----GATTAACGGAA 298
Db 332 TGATATGTTGGATCAATTTTCAAGACAGATCTTAATTTTGAGGCTCACGAGTACATGA 391
QY 299 TGGTGAACCGCTCTTCCTATGTTTCCGCTCTTATCGAAAGCAAGGATATTTATGTTACATG 358
Db 392 TTTTCAAACTTTATCCGTTTCAATTTTCGACTATTGAGACAACTGTTTACAATATCTCCCC 451
QY 359 TGATGTTTCAATTAATTAAGACAAATAATGAGGCTTCAAGCAATCGTTAGCTTAATGA 418
Db 452 AAAACCTTTTATAGATTCGAAGATCCAAAGGCAAAATTTAAAGAAATCTCTTTGTAACGA 511
QY 419 TGTTGAAGGTTTGTGAGTGTGACGAAGCACTTCTATGAGGCTACCTGGGGAGATTAT 478
Db 512 CATCAAGGCTTTTGAACCTTATACGAGCTCGATGTAAAGACTCATGGAGAAATAT 571
QY 479 ATTAGAAGTGTCTTTGGTTTACAGCATCTCGTCTTAGCATTTATGACAAAAGATGCTTT 538
Db 572 TTTGGAAGGCGACTTGTCTTCTACTGCTCATCTTGAATCT-----CGAGCTCC 622
QY 539 TTTTCAAAACCGGCTCTTTTACCGAAATPACAAACGGGCACTAAAGCAACCCCTTTGGAA 598
Db 623 ACATTTGAAGTCAACCTCTGAGTAAGCAAGTGACACATGCGCTTGAGCAATCTCTCCATA 682
QY 599 AAGGTTGCCAAGATAGAGCGCGCGAGTACATCTCTTCTATCAACAAAGATCTTCA 658
Db 683 GAGCATTCGAAGTTGAGACAGCTTCTATCTATCTAGCAAGAGAGAAACAGAA 742
QY 659 TAACAAGACTTTTACTTAACTTGTCTAAGTTAGAGTTCAATTTGCTTCACTTCACTGACAA 718
Db 743 GAATGATGTTGCTTCAATTTGCAAACTGGACTTCACTTACTTCACTTCACTTCACTTCA 802
QY 719 GGAAGACTCAGCATGTGTGCAATGTTGGAAAGCTTTGATATCAAGAAAGCGGAC 778
Db 803 ACAAGAACTTAGTGAAGTATCAAGGTGGTGGAAAGATTTGGATTTGTGCAACACTTCC 862
QY 779 TTGTTTAAAGATAGATTTGTTGAAGTCTATTTTGGGAGTCTAGGTTTCAAGCTTATGAGCC 838
Db 863 ATATGCTAGGATAGAGCAAGTGAATGCTATTTTGAACGATGGGGTGTATGCTGAACC 922
QY 839 ACAGTATTCGGGGCTAGAGTTTCTTCAAAAGCTTTGCTTATTAATCTTATAGA 898
Db 923 TCAATACTCTCAGGCTCGTGTATGCTTGTCTAAGACTATAGCAATGATTTCTATAGTAGA 982
QY 899 TGACACTTATGATGCGTATGTTACTTATGAGAACTTAAGATCTTTTACTGAAAGCTGTTGA 958
Db 983 TGACACATTCGATGCTTATGCTATGCTTCTCAAGAACTTGAATCTACACCGATGCCATACA 1042
QY 959 AAGGTTGCTCAATTTACATGCTTTAGACACACTTCCAGAAATACATGAAACCGATATACAAAT 1018
Db 1043 GAGTGGGATATTTAGCAAAATTTGATCGGCTCCCTGATTTATCATGAATATCGTTTACAAAGC 1102
QY 1019 ATTCATGATACATACAGAAATGGAAGAAATTTCTTGAAGAGGAGGAAAGACAGATCT 1078
Db 1103 ACTTTTATGATCTCAATGATTTATGAATGAGTGTGTCACAGGATGTTAGATCTGATGT 1162
QY 1079 ATTTTAACTCGGCAAGAAATTTGTGAAAGTTTGTGAAACCTGATGTTTGAAGCAAA 1138
Db 1163 TGTTCACTACGCGAAAGAAAGATGAAAGAAATCGTGAGAAATATTTTGTGGAAGCAAA 1222

QY 1139 ATGGGCAATGAGGACACATACCAACCACTGAAGAGCATGATCCAGTTGTTATCATTTAC 1198
Db 1223 ATGTTTCATTGGAAGATATATGCCCCAGTCTCTGAGTATCTTAGCAATGCAATTTAGCTAC 1282
QY 1199 TGGCGGTGCTAAACCTGCTTACAAACAACTTGTATCTTGGCATGAGTATATTTCAAAA 1258
Db 1283 CAGCACTTATTTACTTGTTCAGCACTACATCTTATTTGGCATGAAAGTCTGCTAAACAAGA 1342
QY 1259 AGAGTCTGTGGAATGGGCTGTCTGCACTCTCTCTTTTAGATATCTCAGGTATATCTGG 1318
Db 1343 AGATTTTG---AATGGTTGGCCAAAGAACCTTAAATTTCTTGAGGCTAATGTGACGTTATG 1399
QY 1319 TCGAGGCTTAATGATCTCATGACCCCAAGCGCAGCAAGAAAGAAACATAGTTTCATC 1378
Db 1400 CCGAGTCATGATGACATAGCCACTATGAGGTTGAGAAAGGTAGAGTCAAGTCTCCAC 1459
QY 1379 GAGCTTTGAAAGTTATATGAAGGAATATAATGTCAATGAGGAGTATGCCCAACCTTGAT 1438
Db 1460 TGAATTTGAATGTTTACATGAGAGATTATGGTGTATCCACAGAAAAGGCCATCGAAAAAT 1519
QY 1439 TTCAAGGAAGTAGAAGATGTTGGAAGAGATATAAACCGAGAGTACCTCACAACATAAAA 1498
Db 1520 CCAAGAAATGGCTGAGACAGCATGGAAGGATTAATGAAGGAATCCTTCGACCAACTCC 1579
QY 1499 CATTCGAAGCGCTTATTTGATGGCTGTGATCTATTTGTGCCAGTTTCTTGAAGTTCAATA 1558
Db 1580 CGTCTCTACAGAGATCTCTACTCGCATCTCAATCTTCTGCTCGCATTTATCGATTTACTTA 1639
QY 1559 TGCAGGAA---AGGATAACTTTCACAGTATGCGGAGACGAATACAAACATCTCATAAAGTC 1615
Db 1640 TAAGCACAATCAAGATGATACATCTCATCGGAAAAAGTACTTAAAAACCTCATATTATTGC 1699
QY 1516 TCTACTCGTTATCTCTATGATATGA 1643
Db 1700 GTTGTGCTGACTCTATTGAAATTTAA 1727

RESULT 12
US-09-900-797-31
; Sequence 31, Application US/09900797
; Patent No. 6645762
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 6645762, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/900,797
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 1944
; TYPE: DNA
; ORGANISM: Solanum tuberosum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (57)...(1724)
; OTHER INFORMATION: vetispiradiene synthase
US-09-900-797-31

Query Match 20.0%; Score 330.4; DB 4; Length 1944;
Best Local Similarity 52.0%; Pred. No. 1.5e-87;
Matches 857; Conservative 0; Mismatches 761; Indels 30; Gaps 4;
QY 14 AGAAGAAAAACCTATTGCCCCCAATTGCCAACTTTCTCTCAAGCAATTTGGGAGATCAGTT 73

Db 92 AGAGGAGGAGATTGTTCCGCCCATAGCTGACTTCTCTCAAGTCTTTGGGGTGTGCTTT 151
QY 74 TCTCATCTATCAAAAGCAAGTAGAGCAAGGGGTGGAACAGATAGTGAATGATTTAAAAA 133
Db 152 TCATTTCAATTCCTCCCTCGACAATCAATTCGTGGAATATATGCTCAAGAGATCGAAACTTT 211
QY 134 AGAAGTGGGCAACTACTATAAGAGCTTTGGATATTCCTATGAACATGCCAATTTGTT 193
Db 212 GAAGGAACAATCAAGAAATTAATTAATCTGCAATCTTCTCGAAGCAATTTGGCTTGAGAAAT 271
QY 194 GAAGCTGATGATGAAATTTCAACGCCCTTTGGAATACCGTATCACTTTGAACGGGAGATTGA 253
Db 272 GGATCTGATAGACATTTGTTGAGCGCTTTGGCATTGCTTATCATTTTGAAGAAACAATAGA 331
QY 254 TCATGCAATGCAATGATTAATTAAGAAATATGTTGTT-----GATACTGGAA 298
Db 332 TGATATGTTGGATCAATTTTCAAGAGCAGATCCTAACTTTGAGGCTCAGAGTACAATGA 391
QY 299 TGGTGACCGCTCTTCTTATGTTGTTCCGCTTATGCGAAGCAAGGATATTAATTTACATG 358
Db 392 TTTACAAACTTTATCCGTTCAATTTTGGACTATTTGAGACAACATGTTTACAATATCTCCCC 451
QY 359 TGATGTTTCAATTAATATAAGACAAATAATGGAGCGTTTCAAGCAATCGTTAGCTAATGA 418
Db 452 AAAACTTTTATTAGATTTCCAGATGCAAAAGGCAAAATTTAAGAAATCTCTTTGTAACA 511
QY 419 TGTTGAAGTTTGGCTTGTGATTTGAGAGCAACCTTCTATGAGGGTACCTGGGAGATTAAT 478
Db 512 CATCAAGGGTCTTTGAACTTATAGAGCCCTCGCATGTAAGGACTCATGAGAGAGATAT 571
QY 479 ATTAGAGAGTCTCTTGGTTTTACAGATCTGCTTAGCATTAATGACAAAGATGCTTT 538
Db 572 TTTGGAAGAGCACTTGTCTTCTACTGCTCAATCTTGAATCT-----GCAGCTCC 622
QY 539 TTCTACAAACCCCTCTTTTACCGAAATACAGGGCAGCTTAAAGCAACCCCTTTGGAA 598
Db 623 ACATTTGAGTCACTCTGAGTAAAGCAAGTACATGCGCTTGAAGCAATCTCTCCATTA 682
QY 599 AAGTTGCCAAGATAGAGGGGGCGAGTCAATCTCTTCTATCAACAAAGATCTCA 658
Db 683 GAGCATTTCCAGAGTTGAGACACGCTACTTCACTCTATCTACGAGAGGAGGAGCAAGAA 742
QY 659 TACAAGACTTTACTTAACTTGTAGTTAGATTTCAATTTGCTTCACTAGTCAATTCACAA 718
Db 743 GAATGATGTTGCTTCAATTTGAAAATCGACTTCACTTCACTTCACTAGATTTGTCACAA 802
QY 719 GGAAGAGCTCAGCCATGTTGCAATGTTGGAAGCTTTTCGATATCAAGAAAGCGCAC 778
Db 803 ACAAGAACTTAGTGAAGTATCAAGGTGTTGGAAGATTTGGATTTTGACAACTTCC 862
QY 779 TTGTTAAGAGATAGAAATTTGTAATGCTACTTTTGGGAGCTAGGTTTCAAGGCTATGAGCC 838
Db 863 ATATGCTAGGATAGAGCAGTGAATGCTACTTTTGGACGATGGGGGTGTATGCTGAACC 922
QY 839 ACAGTATTCGGGCTAGAGTTTCTTCAAAAGCTTGTCTGTTATTAACCTTTATAGA 898
Db 923 TCAATCTCTCAGGCTCGTGCATGCTTGTGAAGACTATAGCAATGATTTCTATAGTAGA 982
QY 899 TGACACTTATGATCGGTATGCTATTAAGAACTTAAAGATCTTTACTGAACTGTTGA 958
Db 983 TGACAACTTGAUGCTTATGCAATTTGCAAGAACTTGAAGATCTACCGATGCCATACA 1042
QY 959 AAGTGTGCAATTAATGCTTTAGACACACTTTCCAGAAATACATGAACCCGATATACAAAT 1018
Db 1043 GAGGTGGGATATTAGCCAAATTTGATCGGCTCCCTGATTACATGAATAATCAGTTACAAAGC 1102
QY 1019 ATTCAATGATACATACACAGAAATGGAAGAAATTTCTTCAAGGAGGAGGAGACAGATCT 1078
Db 1103 ACTTTTATGATCTCTACAAATGATTAATGAATGGAGTTGTCCAAAGGATGGTAGATCTGATGT 1162
QY 1079 ATTTAACTGGCGCAAGAAATTTGTGAAGAGTTTGTGTAGAAACCTGATGTTGAAGCAAA 1138

Db 1163 TGTTCACTACGCGAAAGAAAGAAATCGTGTGAGAAACTATTTTGTGGAAGCAAA 1222
QY 1139 ATGGGCAATAGAGGACACATATCAACCACTGAAGAGCATGATCCAGTTTGTAAATCATTTAC 1198
Db 1223 ATGTTTCAATTTGAAGGATATATGCCCGCAGTCTCTGTAGTATCTTTAGCAATGCAATTAGCTAC 1282
QY 1199 TGCGCGTGTCTAACTGCTTACAACAACCTTGTATCTTTGGCATGAGTGATATATTTCAAAA 1258
Db 1283 CAGCACTTAATTAATTTGCTTACGACTACATCTTATTTGGGCATGAAGTCTGCTAAACAGCA 1342
QY 1259 AGAGTCTGTGAATGGGCTGTCTGTGCACCTCTCTTTTATAGATATCTCAGGTATACTTGG 1318
Db 1343 AGATTTTG--AATGGTTGGCCAAAGAACCTTAAAAATCTTTGAGGCTAATGTGACGTTATG 1399
QY 1319 TGGACCCCTTAATGATCTCATGACCCCAAGGCGGACCAAGAAAGAAACATAGTTTCATC 1378
Db 1400 CCGAGTCATAGATGACATAGCCACCTATGAGGTTGAGAAGGTTAGAGGTCAAGTTGCCAC 1459
QY 1379 GAGCCTTTGAAGTTATATGAAGGAATATAATGTCATGAGGAGTATGCCCAACCTTGAT 1438
Db 1460 TGGAAATGAAATGTTATCATGAGAGATTTATGTTATCCACAGAAAGGCCATGGAAAAAT 1519
QY 1439 TTACAAGGAATGAGAAGATGTGTGGAAGATATAAAACCGAGAGTACCTCACAACTTAAAA 1498
Db 1520 CCAAGAAATGGCTGAGACAGCATGGAAGGATGTAATGTAAGGAATCTTTCGACCAACTCC 1579
QY 1499 CATTCGAAGCCGTTATTTGATGCTGTGATCTATTTGTGCCAGTTTCTTTGAAGTTCAATA 1558
Db 1580 CGTCTCTACAGAGATTTCTCACTCGCATTTCTCAATCTTTGCTGCGATTAFCGATGTTACTTA 1639
QY 1559 TGCAGGAA--AGGATAACTTCAACGTATGGGAGACGAATACAAACATCTCTATAAAGTC 1615
Db 1640 TAAGCAATCAAGATGGATACACTATCCGAAAGAAAGTACTATAAACTCATATTTATTGCG 1699
QY 1616 TCTACTCGTTTATTCCTATGAGTATATGA 1643
Db 1700 GTTGTGGTGGACTCTATTGAAATTTAA 1727

RESULT 13
US-09-398-395A-1
; Sequence 1, Application US/09398395A
; Patent No. 6468772
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 64687721, Joseph P.
; APPLICANT: Starke, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/398,395A
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1671
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (25)....(1668)
US-09-398-395A-1

Query Match 19.4%; Score 320.6; DB 4; Length 1671;
Best Local Similarity 52.0%; Pred. No. 1.1e-84;
Matches 850; Conservative 0; Mismatches 764; Indels 21; Gaps 5;

1120	ATAGAAAGATGAAGAAGTAGTAAGAAATTAATGTGCGTCAACATGGTTATTGAA	1179
1152	GGACACATACCAACCACTGAAGAGCATGATCCAGTTGTTAATCATTTACTGGCGGTCTAAC	1211
1180	GGATATTACGCCACCTGTTTCTGTAATACCTAAGCAATGCACCTAGCAATACCACATATTAC	1239
1212	CTGCTTACACAACTGTTTATCTTGGCATGAGTGATATATTCACAAAAGAGTCTGTGAA	1271
1240	TACGTCGCAACATCGTATTTGGGCATGAAGTCTGC--CAGCGACGAAGATTTTGAG	1296
1272	TGGGCTGCTCTGCACCTCCTCTTTTATAGTACTCAGGTATACCTTGGTCGAGCGCTAAAT	1331
1297	TGGTTGTCAAGAATCCAAAATTCCTTGAAGCTAGTGAATTAATGTGCGAGTTATCGAT	1356
1332	GATCTCATGACCCCAAGCGCGAGCAAGAAAGAAAACATAGTTTCATCGAGCCTTTGAAAGT	1391
1357	GACACGCCAGTACGAGTTGAGAAAAGCAGGGGACAAATTGCACTGGAAATTTGAGTGC	1416
1392	TATATGAAGGAATATAATGTCAATGAGGAGTATGCCCAAACCTTGATTTTCAAGGAAGTA	1451
1417	TGCTATGAGAGTATATGTTATATCAACAAGAGCGAATGGCTAAATTTCAAAATATGSGCT	1476
1452	GAAGATGTGTGGAAGATATAAACCGGAGAGTACCTCAACAATAAAACATTTCCAAGGCGG	1511
1477	GAGACAGCATGGAAGATATTAATGAAAGGACTTCTTAGGCCCACTCCCGTCTCTACAGAA	1536
1512	TTATTGTGGCGTGATCTATTGCTGCCAGTTTCTTGAAGTTCAATATGCAGGAAA--G	1568
1537	TTTTTAACCTCATTTCAATCTTGCTGTTATTTGTTGAGGTTACATATACACAATCTA	1596
1569	GATAACTTCAACGTTATGGGAGAGCAATACAAAACATCTCATAAAGTCTCTACTCGTTTAT	1628
1597	GATGGATACACTCATCGGAGAAAAGTCTTAAAACTCACATTTATTAACTACTTGTGGAC	1656
1629	CCTATGAGTATATGA	1643
1657	TCCATCAAAATTTGA	1671

```

RESULT 14
US-09-887-586A-1
/ Sequence 1, Application US/09887586A
/ Patent No. 6495354
/ GENERAL INFORMATION:
/ APPLICANT: Chappell, Joseph
/ APPLICANT: No. 64953541, Joseph P.
/ APPLICANT: Sparks, Courtney M.
/ APPLICANT: Manna, Kathleen R.
/ TITLE OF INVENTION: SYNTHASES
/ FILE REFERENCE: 07678-025001
/ CURRENT APPLICATION NUMBER: US/09/887,586A
/ CURRENT FILING DATE: 2003-06-22
/ PRIOR APPLICATION NUMBER: 09/398,395
/ PRIOR FILING DATE: 1999-09-17
/ PRIOR APPLICATION NUMBER: 60/130,628
/ PRIOR FILING DATE: 1999-04-22
/ PRIOR APPLICATION NUMBER: 60/150,262
/ PRIOR FILING DATE: 1999-08-23
/ NUMBER OF SEQ ID NOS: 58
/ SOFTWARE: FASTSEQ for Windows Version 3.0
/ SEQ ID NO 1
/ LENGTH: 1671
/ TYPE: DNA
/ ORGANISM: Nicotiana tabacum
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (25)...(1668)
US-09-887-586A-1

```

Query Match 19.4%; Score 320.6; DB 4; Length 1671;
Best Local Similarity 52.0%; Pred. No. 1.le-84;
Matches 850; Conservative 0; Mismatches 764; Indels 21;

15 GAAGAAACCTATTTGCGCCCAATTCGCCAATTCCTCCAAAGCATTTGGGGAGATCAGTTT 74
 Db
 52 GAAGAAGAGATTGTTGGCCCGTCGCCGACTTCTCCCTAGTTTGTGGGGTGCATCAGTTC 111
 Qy
 75 CTCATCTATCAAAAGCAAGTATGAGCAAGGGTGGAAACAGATAGTGAATGATTTAAABAA 134
 Db
 112 CTTTCATCTCCATTAA---AAATCAGGTTCAGAAAGATATGCTCAAGAGATTGAAGCA 168
 Qy
 135 GAAGTGGGGCACTACTAAAGAGACCTTTGGATATTCCTTATGAACATGCCAATTTGTTG 194
 Db
 169 TTGAAGGAACAAACGAGGAATATGCTTTAGCAACTGGAATGAAATTTGCTCATACACTG 228
 Qy
 195 AAGCTGATGATGAATCAAGCCCTTGAATACCGTATCACTTTGAACGGGAGATTGAT 254
 Db
 229 AATTGTAGACACTAATTGAACGCTTGGCATATCTTACCACTTTGAGAAAGAAATTTGAT 288
 Qy
 255 CATGCATTTGCAATGATTTATGAACATATGCTGATTAACCTGGAATGGTGACCGCTCTCC 314
 Db
 289 GATATTTTGGATCAGATTTCACCAAACTCAAACTGCAACGATTTGTGCACTTCTGCA 348
 Qy
 315 TTATGGTTCCTTTATGCGAAAGCAAGATATTTATGTTACATGATGATGTTTTCAATAAC 374
 Db
 349 CTTCAAATTTGCTCAGGCAACATGTTTCAACATCTCTCCTGAAATTTTCAGCAA 408
 Qy
 375 TATAAGACAAATATGAGCGCTTCAAGCAATGTTAGCTATGATGTTGAAGTTTGTCTT 434
 Db
 409 TTCCAAGACGAAATGCGAAATTCAGGAATCTCTTGTAGTGTCTTAGGATTTATG 468
 Qy
 435 GAGTTGTACGAAGCAACTTCTATAGGCTACCTGGGAGATATATAGAGATGCTCTT 494
 Db
 469 AACTTGTATGAGCTTCACTGTAAGGACTCATGCTGAGATATCTTAGAAGCGACAT 528
 Qy
 495 GGTTTTACAGATCTGCTTATGCAATATGCAAAAGATGCTTTTTTCTACAAACCCCGCT 554
 Db
 529 GCTTTCTC-----CACTATCCATCTTGAATCTGCAAGGCTCCACATTTGAAATCTCCA 579
 Qy
 555 CTTTTCACGAAATACACGGGCACTAAAGCAACCCCTTTGGAAGGTTGCAAGATTA 614
 Db
 580 CTTAGGAGCAAGTACACATGCTTGGCAATGTTTGCACAGGGTGTCTCTAGAGTC 639
 Qy
 615 GAGCGGGCAGTACATTCCTT---TCTATCAACCAAGATTCCTCAACCAAGCTTTA 671
 Db
 640 GAGACCGATTTCTCATCTCATCTATGCAAGCAACATCGAAGATATGTTGTTA 699
 Qy
 672 CTTAACTTGTAAAGTATGATTCATTTGCTTCACTATTCGCAAGGAAGAGCTCAGC 731
 Db
 700 CTTGATTTGCCAATTTGAATTTCACTTGTCTCCAGATGTTGCAACCAAGACTTGTCT 759
 Qy
 732 CATGTGCAATGTTGGAAGCTTTCGATATCAAGAGAACGACCTTTGTTTAAGAT 791
 Db
 760 CAAGTATCAAGGTGGTGAAGATTTGATTTTGTAAACCACTTCCATATGCTAGAT 819
 Qy
 792 AGAATTTGATGCTACTTTTGGGAGCTAGGTTCAAGGCTATGAGCCAGATTTCCCGG 851
 Db
 820 CGAGTATGATGCTACTTTTGGGCAATGAGAGTTTATTTGAGCTCAATCTCTCAA 879
 Qy
 852 GCTAGAGTTTCTTCAAAAAGCTTGTGTTTATTAACCTTTATAGATGACACTTATGAT 911
 Db
 880 GCTCGGCTCATGCTGTTAAGACCATATCAATGATTTGATTTGATGACACCTTTGAT 939
 Qy
 912 CGGTATGCTATTAAGAACTTAAGATCTTTTACGTAAGCTTGTGAAGGTTGTCATTT 971
 Db
 940 GCTTACGCTACAGTTTAAAGAACTTGAGGCATACACAGATGCCATACAAAGATGGGATC 999
 Qy
 972 ACATGCTTAGACACTTCCAGATACATGAACCGATATACAAATTTATTCATGGATACA 1031
 Db
 1000 AACGAAATGATCGGCTTCTGATTTACATGAAGAAATCAGTTTCAAGCTATTTAGATCTC 1059
 Qy
 1032 TACACAGAAATGGAAGATTTCTTTCAGAGGAGGGAAGAACAGATCTATTTAACTCGGG 1091
 Db
 1060 TACAAGGATTTAAGAAAGAAATTTGCTAGTGGCGGAGATCTCATATTTCTGCGCATGCA 1119

1092 AAGAATTTGTGAAGAGTTTGTAGAAACCTGATGTTGAGCAAAATGGCAATGAG 1151
 Db
 1120 ATAGAAGAAGATGAAGAAGTAGTAAGAATTTATATGTCGAGTCAACATGTTTATTGAA 1179
 Qy
 1152 GGACACATACCAACCACTGAAGAGCATGATCCAGTTGTAATCATTTACTGGGGTCTAAC 1211
 Db
 1180 GGATATACGCCACCTGTTTCTGATACCTAAGCAATGCATGCAACTACACATATTAC 1239
 Qy
 1212 CTGCTTACAACTATGTTTATCTTTGTCATGATGATATTTCAAAAAGAGTCTCTCGAA 1271
 Db
 1240 TACCTCGGCAACATCTGTTATTTGGGCATGAAGTCTGC---CACGGAGCAAGATTTGAG 1296
 Qy
 1272 TGGGCTGCTCTGCACTCTCTTTTATAGATCTCAGGTATCTTGGTCGAGCCCTAAT 1331
 Db
 1297 TGGTTGTCAAAGATCCAAATTTCTTGAAGCTAGTGTATTTATGTCGAGTTATCGAT 1356
 Qy
 1332 GATCTCATGACCCACCAAGGCGGAGCAAGAAAGAAACATAGTTTCATCGAGCTTCAAAAGT 1391
 Db
 1357 GACACAGCCAGCTACGAGGTTGAGAAAGCAGGGGACAAATTTGCACTGGAATTCAGTGC 1416
 Qy
 1392 TATATGAAGAAATATATGTCATGAGGATATGCCCAACCTTGATTTTCAAGGAAGTA 1451
 Db
 1417 TGCATGAGATTTATGTTATATCAACAAAGAGGCAATGCTTAAATTTCAAAATATGGCT 1476
 Qy
 1452 GAGATGCTGGAAGATATTAACCGGAGATCTCACTCACTAAACATTTCCAGGCGG 1511
 Db
 1477 GAGACGATGGAAGATATTAATGAAGGACTTCTTAGGCCACTCCCGTCTCTACAGAA 1536
 Qy
 1512 TTATTCATGCTGCTGATCTATTTGTGCCAGTTTCTTGAAGTTCAATATGCAAGAAA---G 1568
 Db
 1537 TTTTAACTCTTATCTCATCTTGTCTGATTTGTTGAGGTTACATATACACATCTA 1596
 Qy
 1569 GATACTTACAGTATGAGGAGCAAGATCAAACTCTATAAAGTCTCTACTCTGTTAT 1628
 Db
 1597 GATGGATACACTCATCGGAGAAAGTCTTAAACCTCACATTTATAOCTTACTTGTGGAC 1656
 Qy
 1629 CCTATGATATATGA 1643
 Db
 1657 TCCATCAAAATTTGA 1671

RESULT 15

US-09-895-752-1
 ; Sequence 1, Application US/09895752
 ; Patent No. 6559297
 ; GENERAL INFORMATION:
 ; APPLICANT: Chappell, Joseph P.
 ; APPLICANT: No. 6559297, Joseph P.
 ; APPLICANT: Starks, Courtney M.
 ; APPLICANT: Manna, Kathleen R.
 ; TITLE OF INVENTION: SYNTHASES
 ; FILE REFERENCE: 07678-025001
 ; CURRENT APPLICATION NUMBER: US/09/895,752
 ; PRIOR FILING DATE: 2001-06-29
 ; PRIOR APPLICATION NUMBER: 09/398,395
 ; PRIOR FILING DATE: 1999-09-17
 ; PRIOR APPLICATION NUMBER: 60/100,993
 ; PRIOR FILING DATE: 1998-09-18
 ; PRIOR APPLICATION NUMBER: 60/130,628
 ; PRIOR FILING DATE: 1999-04-22
 ; PRIOR APPLICATION NUMBER: 60/150,262
 ; PRIOR FILING DATE: 1999-08-23
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 1
 ; LENGTH: 1671
 ; TYPE: DNA
 ; ORGANISM: Nicotiana tabacum
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (25)...(1668)
 US-09-895-752-1

Query Match 19.4%; Score 320.6; DB 4; Length 1671;
Best Local Similarity 52.0%; Pred. No. 1.le-84;
Matches 850; Conservative 0; Mismatches 764; Indels 21; Gaps 5;

QY 15 GAAGAAACCTATTGCGCCCAATTCGCCAACTTCTCCCAAGCATTTGGGAGATCAGTTT 74
DB 52 GAAGAGAGATTGTTGCGCCCGTCCGCCACTTCTCCCTAGTTTGTGGGGTGCATCAGTTC 111
QY 75 CTCATCTATCAAAAGCAAGTAGAGCAAGGGTGGAAACAGATAGTGAATATTAAAAA 134
DB 112 CTTTCATTCTCAATTA--AAATCAGGTTGAGAAAGATAGTCTCAAGAGATTGAAGCA 168
QY 135 GAAGTGGCGCACTACTAAAAGAACTTTGGATATTCTCTATGAACATGCCAATTGTTG 194
DB 169 TTGAAGGAACAAACGAGGAATATGTTGTAGCAACTGGAATGAAATGSCGTGATACATG 228
QY 195 AAGCTGATTGATGAATTCACAGCCTTGGAAATACCGTATACCTTTGAACGGGAGATTGAT 254
DB 229 AATTGTAGACACTATTGAACGCTTGGCATATCTACCACTTTGAGAAAGAAATTGAT 288
QY 255 CATGCATTGCAATGTATTATTGAACATATGTTGATTAACCTGGAATGGTGCCGCTCTCC 314
DB 289 GATATTTGGATCAGATTTCACCAAACTCAAACTGCAACGATTTGTGCACTTCTGCA 348
QY 315 TTATGGTTCGTTCTATGCGAAGCAAGGATATTATGTTACATGATGATGTTTCAATAAC 374
DB 349 CTTCAATTTGATTCGTCAGGCAACATGGTTTCAACATCTCTCCCTGAATTTTCAGCAA 408
QY 375 TATAAGACAAAATGAGCGGTTCAAGCAATCGTTAGCTAATGATGTTGAAGGTTTGCTT 434
DB 409 TTCCAAGACGAAATGGCAAAATTCAGGAATCTTGTCTAGTGTGCTTTAGGATTATTG 468
QY 435 GAGTTGTACGAAGCAACTTCTATGAGGTTACCTGGGAGATTATATTAGAAGATGCTCTT 494
DB 469 AACTTGATGAGCTTCAATGTAAGGACTCATGCTGAGATATCTTTAGAAGCGCACTT 528
QY 495 GGTTTTACAGCATCTGCTTTAGCATTTATGACAAAGATGCTTTTCTACAAACCCCGCT 554
DB 529 GCTTTCTC-----CACTATCCATCTTGAATCTGCAAGCTCCACATTTGAAATCTCCA 579
QY 555 CTTTTTACCGAAATACACGGGCACTAAGCAACCCCTTTGGAAAGGTTGCGCAAGATA 614
DB 580 CTTAGGAGCAAGTGACACATGCCCCCTTGAGCAATGTTTGCAACAGGGTGTCTCTAGATC 639
QY 615 GAGCGGCGCAGTACATTCCTT---TCTATCAACAAGATTCATCAACAAGACTTTA 671
DB 640 GAGACCGGATCTTCACTCTCATCATCTATGACAGGAAACATCGAAGATTAATGTGTGA 699
QY 672 CTTAAACTTGCTAAGTTAGATTCAATTTGCTTCAGTCAATTCGCAACAGGAGAGCTCAGC 731
DB 700 CTTGATTTGCCAAATTTGAATTTCAACTTGTCTCCAGATGTTGCAACAAACAGAACTTGT 759
QY 732 CATGTTGCCAAATGGTGGAAAGCTTTCGATATCAAGAAAGACGCACTTGTGTTAAGAT 791
DB 760 CAAGTATCAAGGTGGTGGAAAGATTGGATTTTGTAAACAACACTTCCATATGCTAGAT 819
QY 792 AGAATTGTTGAATGCTACTTTTGGGAGCTTAGTTCAGGCTATGAGCCACAGTATCCCGG 851
DB 820 CGAGTAGTTGAATGCTACTTTTGGGCATTAGAGTTTATTTGAGCCTCAATFACTCTCAA 879
QY 852 GCTAGAGTTTCTTCAAAAAGCTTGTGTTGTTATAACTCTTATAGATGACACTTATGAT 911
DB 880 GCTCGGTCATGCTCGTTTAAGACCATATCAATGATTTGATTTGATGATGACACCTTTGAT 939
QY 912 GCGTATGTTGTTATGAGAACTTAAAGATCTTTACTGAAGCTGTTGAAAGGTTGGTCAATT 971
DB 940 GCTTACGGTACAGTTAAAGAACTTGAGGCATATGAGAGTTTATTTGAGCCTCAATFACTCT 999
QY 972 ACATGCTTAGACACACTTCCAGAAATACATGAACCGATATACAAATTTATTCATGGATACA 1031
DB 1000 AACGAAATTGATCGGCTTCTGATTACATGAAATACAGTTTACAAAGCTATTCTAGATCTC 1059
QY 1032 TACACAGAAATGGAGAAATTTCTTGCAAGGAGGGAAGAACAGATCTATTAACTCGCGC 1091

DB 1060 TACAAGGATTATGAAAGGAATTTGCTTAGTGCCGAGAGATCTCATATTGTCTGCCATGCA 1119
QY 1092 AAGAATTTGTGAAAGAGTTTGTAGAAACCTGTAGTTGTGAGGAAATTTGGGCAATAGAG 1151
DB 1120 ATAGAAAGATGAAAGAGAGTAGTAAGAAATTTATAATGTCGAGTCAACATGTGTTATTGAA 1179
QY 1152 GGACACATACCAACCACTGAGAGCATGATCCAGTTGTAAATCATTTACTGGGGTGTCTAAC 1211
DB 1180 GGATATACGCCACTGTTTCTGAAATCTTAAGCAATGCACTAGCAACTACCAATATTAC 1239
QY 1212 CTCCTTACAACTTTGTTATCTTGGCATGAGTGATATATTCACAAAAAGAGTCTCTCGAA 1271
DB 1240 TACCTCGGACAAACATCGTATTTGGGCATGAAGTCTGC--CACGGAGCAAGATTTTGAG 1296
QY 1272 TGGGCTGTCTCTGCACCTCTCTTTTAGATATCTCAGGTATACTTGTGCGAGCCCTAAAT 1331
DB 1297 TGGTTGTCAAGAAATCCAAAATTTCTTGAAGCTAGTGTAAATTATATGTGCGATTATCGAT 1356
QY 1332 GATCTCATGACCCACAAAGCGGAGCAAGAAAGAAACATAGTTTCATCGAGCTTCAAAGT 1391
DB 1357 GACACAGCCACGTACGAGGTTGAGAAAGCAGGGGACAAATTGCAACTGGAAATTGAGTGC 1416
QY 1392 TATATGAAGGAATATTAATGTCAATGAGGAGTATGCCAAACCTTTGATTTACAAGGAATGA 1451
DB 1417 TGCATGAGAGATTATGGTATATCAACAAAGAGGCAATGGCTAAATTTCAAAATATGGCT 1476
QY 1452 GAAGATGTGGAAGAAAGTATAAACCAGAGAGTACCTCACACTAAAACATTTCCAGGCGG 1511
DB 1477 GAGACAGCATGGAAGAGATATTAAATGAAGGACTTCTTAGGCCCACTCCCGTCTCTACAGAA 1536
QY 1512 TTATTGATGGCTGTGATCTATTGTCAGTTCCTTGAAGTTTCAATATGCGAGGAAA---G 1568
DB 1537 TTTTAACTCCTTATTCTCAATCTTGTCTGTTGTTGAGGTTACATATATACAAATCTA 1596
QY 1569 GATAACTTCACAGTATGGGAGAGCAATACAAAACATCTCAATAAGTCTCTACTCGTTTAT 1628
DB 1597 GATGGATACACTCATCCGGAGAAAGTCTTAAACCTCACATTAATTAACCTACTTGTGGAC 1656
QY 1629 CCTATGAGTATATGA 1643
DB 1657 TCCATCAAAATTTGA 1671

Search completed: June 7, 2004, 12:41:20
Job time : 159 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 7, 2004, 11:23:35 ; Search time 754 Seconds
(without alignments)
9977.097 Million cell updates/sec

Title: US-09-763-822A-13

Perfect score: 1649

Sequence: 1 cctggcacttacagagaa.....ctatgagtatataggatcc 1649

Scoring table:

IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched:

2995936 seqs, 2280998010 residues

Total number of hits satisfying chosen parameters:

5991872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	1017	61.7	1641	16	US-10-411-066-37
2	509.6	30.9	1644	16	US-10-411-066-39
3	357.6	21.7	2024	9	US-09-887-586A-51
4	357.6	21.7	2024	9	US-09-903-012-51
5	357.6	21.7	2024	10	US-09-900-797-51
6	357.6	21.7	2024	13	US-09-893-820-51
7	351.6	21.3	1937	13	US-10-424-599-43432
8	330.4	20.0	1944	9	US-09-887-586A-31
9	330.4	20.0	1944	9	US-09-903-012-31
10	330.4	20.0	1944	10	US-09-900-797-31
11	330.4	20.0	1944	13	US-09-893-820-31
12	320.6	19.4	1671	9	US-09-887-586A-1
13	320.6	19.4	1671	9	US-09-903-012-1
14	320.6	19.4	1671	10	US-09-900-797-1

15	320.6	19.4	1671	13	US-09-893-820-1	Sequence 1, Appli
16	316.6	19.2	1644	9	US-09-887-586A-7	Sequence 7, Appli
17	316.6	19.2	1644	9	US-09-903-012-7	Sequence 7, Appli
18	316.6	19.2	1644	10	US-09-900-797-7	Sequence 7, Appli
19	316.6	19.2	1644	13	US-09-893-820-7	Sequence 7, Appli
20	315.4	19.1	1644	9	US-09-887-586A-11	Sequence 11, Appli
21	315.4	19.1	1644	9	US-09-903-012-11	Sequence 11, Appli
22	315.4	19.1	1644	10	US-09-900-797-11	Sequence 11, Appli
23	315.4	19.1	1644	13	US-09-893-820-11	Sequence 11, Appli
24	311.8	18.9	1644	9	US-09-887-586A-3	Sequence 3, Appli
25	311.8	18.9	1644	9	US-09-887-586A-5	Sequence 5, Appli
26	311.8	18.9	1644	9	US-09-887-586A-9	Sequence 9, Appli
27	311.8	18.9	1644	9	US-09-903-012-3	Sequence 3, Appli
28	311.8	18.9	1644	9	US-09-903-012-5	Sequence 5, Appli
29	311.8	18.9	1644	9	US-09-903-012-9	Sequence 9, Appli
30	311.8	18.9	1644	10	US-09-900-797-3	Sequence 3, Appli
31	311.8	18.9	1644	10	US-09-900-797-5	Sequence 5, Appli
32	311.8	18.9	1644	10	US-09-900-797-9	Sequence 9, Appli
33	311.8	18.9	1644	13	US-09-893-820-3	Sequence 3, Appli
34	311.8	18.9	1644	13	US-09-893-820-5	Sequence 5, Appli
35	311.8	18.9	1644	13	US-09-893-820-9	Sequence 9, Appli
36	256	15.5	1994	9	US-09-887-586A-41	Sequence 41, Appli
37	256	15.5	1994	9	US-09-903-012-41	Sequence 41, Appli
38	256	15.5	1994	10	US-09-900-797-41	Sequence 41, Appli
39	256	15.5	1994	12	US-10-041-018-366	Sequence 366, App
40	256	15.5	1994	13	US-09-893-820-41	Sequence 41, Appli
41	248	15.0	1938	13	US-10-425-114-5550	Sequence 5550, Ap
42	225.6	13.7	1650	10	US-09-860-282A-4	Sequence 4, Appli
43	225.6	13.7	1650	10	US-09-860-282A-6	Sequence 6, Appli
44	225.6	13.7	1650	10	US-09-860-282A-10	Sequence 10, Appli
45	225.6	13.7	1650	10	US-09-860-282A-12	Sequence 12, Appli

ALIGNMENTS

RESULT 1
US-10-411-066-37
; Sequence 37, Application US/10411066
; Publication No. US20040005678A1
; GENERAL INFORMATION:
; APPLICANT: KEASLING, JAY
; APPLICANT: MARTIN, VINCENT
; APPLICANT: PITERA, DOUGLAS
; APPLICANT: WITHERS III, SYDNOR T.
; APPLICANT: NEWMAN, JACK
; TITLE OF INVENTION: BIOSYNTHESIS OF AMORPHA-4,11-DIENE
; FILE REFERENCE: 2000-0007.20
; CURRENT APPLICATION NUMBER: US/10/411,066
; CURRENT FILING DATE: 2000-04-09
; PRIOR APPLICATION NUMBER: 10/006,909
; PRIOR FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 37
; LENGTH: 1641
; TYPE: DNA
; ORGANISM: Artemisia annua
US-10-411-066-37

Query Match	61.7%	Score 1017;	DB 16;	Length 1641;
Best Local Similarity	76.2%	Pred. No. 2.2e-263;		
Matches 1251;	Conservative 0;	Mismatches 390;	Indels 0;	Gaps 0;
QY	3	ATGGCACTTACAGAGAAACCTATTTCGCCCACTTCGCACTTTCCTTCCCAAGCATTGG	62	
Db	1	ATGGCCCTGACCGAAGAAACCGATCCGCCCGATCGCTAACTTCCCGCGCTCTATCTGG	60	
QY	63	GGAGATCAGTTTCTCATCTATCAAAAGTAGAGCAAGGGGTGGAGACAGATAGTGAAT	122	
Db	61	GGTGACCACTTCTGTGATCTACGAAAGCAGGTGAGCAGGGTGTGACAGATCGTAAC	120	
QY	123	GATTTAAAAAGAGAGTGGCGCACTACTATAAAGAGCTTTGGATATTCCTATGAACAT	182	

Db 121 GACCTGAAGAAAGATTCGTGAGCTGCTGAAAGAGCTCTGGACATCCCGATGAACAC 180
Qy 183 GCCAATTTGTTGAAGCTGATGATGAATTTCAACGCTTGGATACCGTATCACTTTGAA 242
Db 181 GCTAACCTGCTGAACTGATGACGAGATCCAGCGTCTGGGTATCCCGTACCACCTTCGAA 240
Qy 243 CGGAGAGTTGATCATGCTGCAATGTTATTTATGAAACATATGGTGATGAACCTGGAAATGGT 302
Db 241 CGCGAATCGACACGCACTGCAAGTGCTATACGAAACCTACGGCGACAACTGGAAACGGC 300
Qy 303 GACCGCTCTTCTTATGTTCCGTTATGCGCTTATGCGAAAGCAAGATATATGTTTACATGATGAT 362
Db 301 GACCGTTCTTCTGTTGTTTCTGCTGATGCGTAAACACAGGGCTACTACGTTTACCTGTGAC 360
Qy 363 GTTTTCAATACTATAAGACAAATGGAGGCTTCAAGCAATCGTTAGCTAATGATGTT 422
Db 361 GTTTTCAAACTACAAGCAAGAACGGTGTCTTCAACACAGTCTCTGGCTAACGACGTT 420
Qy 423 GAAGGTTTGTGTTGTTGTAAGCAACATCTTATGAGGGTACCTGGGAGATTTATTA 482
Db 421 GAAGGCTGTGGAACCTGACGAAGGACCTCCATGCGGTGACGGGTGAATCATCTCTG 480
Qy 483 GAAGATGCTCTTGGTTTACAGATCTCGTCTTAGCATATGACAAAAGATGCTTTTCT 542
Db 481 GAGGACGGCTGGGTTTCAACCGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Qy 543 ACAAACCGCTCTTTTACCGAATAACAACGGCACTAAACCAACCCCTTTGGAAAGG 602
Db 541 ACTAACCGGCTCTGTTCAACGAAATCAGCGTGTCTGAAACAGCGGCTGTGAAACGT 600
Qy 603 TTGCAAGATAGAGCGGCGAGTACATCTCTTCTATCAACAACAAGATCTCATAC 662
Db 601 CTGCGGCTATCGAAGCAGCAGTACATCTCGTTTACCAGCAGCAGGACTCTCACAC 660
Qy 663 AAGACTTTTAACTGCTTAAGTTAGATTTCAATTTGCTTTCAGTTCATGACAAAGAA 722
Db 661 AAGACCTGCTGAACCTGGCTAAGCTGAAATTTCAACCTGCTGCACTCTGCAACAAGAA 720
Qy 723 GAGCTCAGCATGCTGCAATGTTGGAAGCTTTCGATATCAAGAAGAACCCACTTCT 782
Db 721 GAATGCTCAGTTTGAAGTGGGAGGCAATTTGACATCAAGAAACCGCGCTGC 780
Qy 783 TTAAGAGATAGAAATTTGTTGAATGCTACTTTTGGGACTAGTTTACGGCTATGAGCCACAG 842
Db 781 CTGCGTGACCGTATCGTTGAATGTTACTTCTGGGTCTGGGTCTGGTTATGACACACAG 840
Qy 843 TATTCGGGCTAGAGTTTCTTCAAAAAGCTGTTGCTGTTATTAACCTTTATAGATGAC 902
Db 841 TACTCCCGTGCACGCTGTTCTTCACTAAAGCTGATGTTATCACCCCTGATCGATGAC 900
Qy 903 ACTTATGATGCTGTTACTTATGAAGAACTTGAAGTCTTTTACTGAAGCTGTTGAAGG 962
Db 901 ACTTACGATGTTACGGCACCTTACGAAGAACTGAAGATCTTTACTGAAGCTGTAGAACGC 960
Qy 963 TGGTCAATTAATGCTTAGACACACTTCCAGAAATACATGAACCCGATATACAAATTTATTC 1022
Db 961 TGGTCTATCACTTGGCTGACACTCTGCGGAGTACATGAACCCGATCTACAACTGTTTC 1020
Qy 1023 ATGGATACATACAGAAATGAAAGATTTCTTTCGAAAGAGGAGGAGAAACAGATCTATTT 1082
Db 1021 ATGGATACCTACACCGAAATGAGGAAATCTCTGGCAAAAAGAGCGCTACCGACCTGTTTC 1080
Qy 1083 AACTGCGCAAGATTTGTGAAAGATTTGTTAGAAACCTGATGTTGTTGAAGCAAAATG 1142
Db 1081 AACTGCGGTAAGAGTTGTTTAAAGATTTCTGTAAGTAACTGATGTTGTTGAAGTAAATG 1140
Qy 1143 GCAATAGGGGACATACCAACCTAGAGGATGATCCAGTTGTTGTAATCATTTACTGGC 1202
Db 1141 GCTACGAGGCCATATCCGACTACGAGAACATGACCCCGTTGTTATCATCACCGGC 1200
Qy 1203 GGTGTAACTGCTTAACAACATTTGTTTATCTTGGCATGATGATATATCAAAAGAG 1262

Db 1201 GGTGCAAACTGCTGACCACTTGTCTATCTGGGTATGTCGACATCTTTTACCAAGAA 1260
Qy 1263 TCTGTGCAATGGGCTGCTCTGACACCTCTCTTTTAGATACTCAGGTATATCTTGGTCGA 1322
Db 1261 TCTGTTGAATGGGCTGTTCTGCAACCCGCTGTTCCGTTACTCGGTATTTCTGGGTGCT 1320
Qy 1323 CGCTAAATGATCTCATGACCCCAAGGCGAGCAAGAAAGAAACAATGATTCATCGAGC 1382
Db 1321 CGTCTGAACGACCTGATGACCCCAAGCAGAGCAGGAAACGTAACACACTCTTCTCTCTCT 1380
Qy 1383 CTTGAAAGTTATGAGGAATATATGTCATGAGGAGTATGCCAAACCTTGAATTTAC 1442
Db 1381 CTGGAATCTTACATGAAGGAATATAACGTTAAACGAGAGTACGACACACTCTGATCTAT 1440
Qy 1443 AAGGAAGTAAAGATGTTGCAAGATATAAACCAGAGTACCTCAACAACCTAAAAACAT 1502
Db 1441 AAGAAGTTGAAGACGATGAGAAAGACATCAACGCTGAATCTGACTACTTAAAAACATC 1500
Qy 1503 CCAAGCCGCTTATGATGCTGATCTATTTGTCAGTTTCTTGAAGTTCAATATGCA 1562
Db 1501 CCGCGCCGCTGCTGATGCGAGTAATCTACCTGTGCCAGTTCTCTGGAAGTACAGTACGCT 1560
Qy 1563 GGAAGGATAACTTCAACGATGAGGAGACGATACAAACATCTCATAAAGTCTCTACTC 1622
Db 1561 GSTAAAGATAACTTCACTCGCATGGGCGCAATACAAACACCTGATCAAAATCCCTGCTG 1620
Qy 1623 GTTTATCTCTATGATGATATGA 1643
Db 1621 GTTTACCGGATGTCATCTGA 1641

RESULT 2

US-10-411-066-39
; Sequence 39, Application US/10411066
; Publication No. US20040005678A1
; GENERAL INFORMATION:
; APPLICANT: KEASLING, JAY
; APPLICANT: MARTIN, VINCENT
; APPLICANT: PITERA, DOUGLAS
; APPLICANT: WITHERS III, SYDNOR T.
; APPLICANT: NEWMAN, JACK
; TITLE OF INVENTION: BIOSYNTHESIS OF AMORPHA-4,11-DIENE
; FILE REFERENCE: 2000-0007.20
; CURRENT APPLICATION NUMBER: US/10/411,066
; CURRENT FILING DATE: 2000-04-09
; PRIOR APPLICATION NUMBER: 10/006,909
; PRIOR FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 39
; LENGTH: 1644
; TYPE: DNA
; ORGANISM: Artemisia annua
US-10-411-066-39

Query Match 30.9%; Score 509.6; DB 16; Length 1644;
Best Local Similarity 57.3%; Pred. No. 1.8e-126;
Matches 942; Conservative 0; Mismatches 699; Indels 3; Gaps 1;

Qy 3 ATGGCACTTACAGAAACCACTTATTCGCCCACTTCCCACTTTCTCCAGCATTTGG 62
Db 1 ATGAGCCTGATCTAGAGGCGTAATCCGCCGCAACCTAACTCCCGCTCGAGATCTGG 60
Qy 63 GGAGATCAGTTTCTCATCTCTCAAAAGAGTAGCAAGGGGTGGAACAGATAGTGAAT 122
Db 61 GGTGACCACTTCTGGCATACGACAGGAGGTTGTAACAGGTTATCAAA 120
Qy 123 GATTTAAAAAGAGTGGCGCAACTACTAAAAAGAGCTTTGGATATTTCTTATGAAACAT 182
Db 121 GACCTGAAGAAGAGGTTAAATCTGAGCTGCTGACCGCTGTGAACCTCCCGACTCAGCAC 180
Qy 183 GCAATTTGTTGAGCTGATGATGAATAATCAACGCCCTTGAACACCGTATCACTTTGAA 242

Db 181 ACTGAAGTCTGAAATTCATCGACCGATCGAGCGTCTGGGTATCGCTTACTACTTCGAA 240
Qy 243 CGGAGATTGATCATGCAATGTAATTTATGAACATATGGTGAATPAACCTGGAATGGT 302
Db 241 GAAGAAATCAACAGGATTTTACGACATGTAACCGCTTACCGCGACATAGTGGACCGGT 300
Qy 303 GACCGCTCTCTTATGGTTCGGTCTTATGCGAAGCAAGATATTAATGTTACATGTGAT 362
Db 301 GGCAACATCTCTCTGTGGTTCGGTCTGATGCGTCAAGCGTCTTCTTCGTAAGCTCTGAC 360
Qy 363 GTTTTCAATTAATATAAGACAAAATGGAGCGTTTCAAGCAATCGTTAGCTTAATGATGTT 422
Db 361 ATCTTCTCCACCTACAGGACAGGAGGTGTTTCAAGAAATCTCTGGAGAGACGTT 420
Qy 423 GAAGGTTTGTGAGTTGTACGAAGCAACTTCTATGAGGGTACTTGGGGAGATATATTA 482
Db 421 CACGGCTCTGGAATCTGACGAAGCGCATACATGTTTGTACCGGTGAAGGATCTCTG 480
Qy 483 GAAGATGCTCTGTTTACAGATCTGCTTACGATATGACAAAGATGCTTTTCT 542
Db 481 GACGACGCGTGTCTTCAACCGGTACTTGTGACGAGATCGCTAAGAACCGGTCCTG 540
Qy 543 ACAACCCCGCTCTTTTACCGAATACACCGGCACATAAAGCAACCCCTTGGAAAAGG 602
Db 541 TCTAATCCCGCTGATCTCTCCAGATCGGTGAAGCTGTGACTCAGCGCTGCATPAGCGT 600
Qy 603 TTCCCAAGATAGAGGGCGGCGACATCTCTTCTATCAACAAAGATTTCTCATTAAC 662
Db 601 CTGCGCGTCTGGAAGCACTGCGTTACATCCGTTCTACCAAGCAGCAGGCACTCTCACTCC 660
Qy 663 AAGACTTATTAACCTGTAAGTTAGTGTCAATTTGCTTCACTTCACTTCACTGACAGGAA 722
Db 661 GAGACCTGTGAACTGGCTAAGCTGGTTCACACCGCTGCACTCTCTGCACAGAAA 720
Qy 723 GAGCTCAGCCATGTCGAAATGTGGAAGCTTTCGATATCAAGAAAGCAAGCACTTGT 782
Db 721 GAAGTGTCTATCTTCTAATGTGGAATCTTTCGAGCTGTGATCAACCTGCGGTAC 780
Qy 783 TTAAGAGATAGATGTTGATGCTACTTTTGGGACTAGTTCAGGTATGAGCCACAG 842
Db 781 GCTGTAACCGTCCGTTGAATGCTACTTCTGGGCTCTGGCTGTGCTACTTTGAACCAAG 840
Qy 843 TATTCGGGGTAGAGTTTCTTCAAAAGCTGTTGCTGTATTAACCTCTTATAGATGAC 902
Db 841 TACTCCGAGTCTCGTCTCTCTCTGTTTCTTCTATCCAGACCTTCTCTGATGAC 900
Qy 903 ACTTATGATCGTATGCTATGAGAACTTAAAGCTTTTACTGAAGCTGTTCAAGAGG 962
Db 901 ACTTACGATGTTACGGCACTTACGAAGACTTGAAGAGTTTACTGAAGCTATCCAGCGT 960
Qy 963 TGCTCAATTACATGCTTAGACACACTTCCGAATACATGAACACCGATATCAAAATTTTC 1022
Db 961 TGGTCTATCACTTGCCTGGACGGTCTGCGGAGTCCATGAATGATCTTCCAGATGCTG 1020
Qy 1023 ATGATACATACACAGAAATGGAAGATTTCTTCAAGAGGAGGGAAGACAGATCTATTT 1082
Db 1021 GTAAAGATCTTCGAAGAAATCGAAGAAATCTGTCTAAGAGCGGCAACAGCACCGCTT 1080
Qy 1083 AACTGCGGCAAGATTTGTCAAGAGTTTGTAGAACTTGTAGAACCTGATGGTGAAGCAAAATGG 1142
Db 1081 AACTACATCAAGAGAGTCTTGAAGAGACGATACAGTCTTATGATGAAGCTGTTGG 1140
Qy 1143 GCAATGAGGACACATACCAACCACTGAAGAGCATGATCCAGTTGTAATCAATTAAGTGC 1202
Db 1141 GCTAAGAGAAATATACCCGACTATCGAAGACATACCAAAAGTTTCTTACATCTCCATC 1200
Qy 1203 GGTGCTAACCTGCTTACAACTGTTTATCTTGGCATGATGATATATTCACAAAGAG 1262
Db 1201 GGTACAACTGCGGCTGTTGCTGGCTTGTGTTATGGGCGAGCTTATCCGGATGAC 1260
Qy 1263 TCTGTGCAAGTGGCTGCTGCACTCTCTTTTATAGTACTCAGGTATCTCTGTCGA 1322
Db 1261 TCTTCAATGGGTATTTACCAACCCCTCCGCTGGTGAAGCGTTGCTGCTGTCGCT 1320

Qy 1323 CCGCTAATGATCTCATGACCCACAGGCGGACGAAGAAACATAGTTTCTATCGAGC 1382
Db 1321 ACTATGGACGACCTGGCTCCCAAAAGGCGAGCAGACCGTAAACACGTTGCTTCCACT 1380
Qy 1383 CTGTAAGGTATATGAAGGAATATAATGTCAATGAGAGATGATGCCAAACCTTGAATTTAC 1442
Db 1381 ATCGAATGCTACATGAAGCAGTTTGAAGCTAGCGAGCAGCAGCATAGATCTCTGAAC 1440
Qy 1443 AAGGAAGTAGAAGATGTGTGAAAGATATAAAACCGAGAGTAC---CTCACAACTTAAAC 1499
Db 1441 AAGAAAGTTGAAGACGATGGAAAGAAATCAACCGTGAATTCATGATCACTTGTAAAGAC 1500
Qy 1500 ATTCAAGCCCGCTTATGATGGCTGTGATCTATTTGTGCCAGTTTCTTGAATTCATAT 1559
Db 1501 GTAAACATCATGATGAGCGATGGGTGACTGAACCTTCTCCGCTCCGCTGAGTACTGTAC 1560
Qy 1560 GCAGGAAGGATAAATTTACACGCTATGGGAGAGCAATACAAACATCTCATAAAGTCTCTA 1619
Db 1561 AAGAACAAAGATCACTTCACTACGTTGGTGTAGAGTTATCAACCAATCAATCCCTG 1620
Qy 1620 CTCGTTTATCTTATGATATATGA 1643
Db 1621 TTCGTTGACGCGATCATCCTGA 1644

RESULT 3

US-09-887-586A-51
; Sequence 51, Application US/09887586A
; Patent No. US20020094556A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20020094556A1, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/887,586A
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 51
; LENGTH: 2024
; TYPE: DNA
; ORGANISM: Lycopersicon esculentum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (32)...(1675)
; OTHER INFORMATION: VFNT germaecene C synthase
US-09-887-586A-51

Query Match 21.7%; Score 357.6; DB 9; Length 2024;
Best Local Similarity 52.7%; Pred. No. 2.1e-85;
Matches 861; Conservative 0; Mismatches 749; Indels 24; Gaps 3;

Qy 29 TGCCCCATGCAACTTTCCTCCAGCAATTTGGGAGATCAGTTTCTCATCTATCAAA 88
Db 58 TGCCCCCTTGGCTTAATTTTCCACCCATCTGTTGGGATATCAATTTCTTATCTCA 117
Qy 89 GCAAGTAGACGAGGGGTGGAACAGATAGTGAATTAATAAAGAAAGTCCGCGCACT 148
Db 118 TGAATTAATCAAGAAAAGTTGAGTTGATGAGTGAAGAGCAATTAGAAAAT 177
Qy 149 ACTAAAGAGCTTTTGGATATCTTATGAACATGCCAATTTGTGAAGCTGATGTGA 208
Db 178 GCTGGTGGAACTTTCGACAAATAGCACTCAAAA-----GCTGTGTTGATAGCGC 228

209 AATTCAACGCTTGGATACCGTATCACCTTTGAACGGAGATTGATCATGCTTGAATG 268
229 GATGCAACGATTGGAGTGGCTTATCATTTGATATGAAATTTGAACATCCATTCAAAA 288
269 TATTATGAACAT-----ATGGTGATAACTGGAATGGTGACCGCTTCTTCCTT 316
289 CATTTTGTATGATCGTCCAAACAGAAATGATAATGACAAACCTTTACGTTGTCTCT 348
317 ATGGTTCGCTTATGCGAAGCAAGGATATTATGTTATCATGATGTTTTCATTAACIA 376
349 TCGTTTTCGACTTGTGAGGCAACAGGCCATTACATGTTCTTCAGATGTTGTTCAGCAAT 408
377 TAAAGACAAAATGAGCGCTTCAAGCAATCGTTAGCTAATGATGTTGGAAGTTTGTCTGA 436
409 CACCAACCAAGATGGAATTTCAAGAAACACTTACTAATGATGTTCCAGGATATTGAG 468
437 GTTGTACGAAGCAACTTCTATGAGGCTACCTCGGGAGATTATATTAGAAGATCTCTTGG 496
469 TTTGTATGAAGCATCACATCTGAGAGTGGCTAATGAGGAGATCTTCTTGAAGAGCTCTTAC 528
497 TTTTACAGATCTGCTGTAGCATTTATGACAAAGATGCTTTTCTTACAAACCCGCTCT 556
529 ATTTACCACTCATCTCGAGTCTTATGCTTCTTCAACTTGAAGCAATTAATAAATCTCT 588
557 TTTTACCGAATACAAACGGGCACTAAAGCAACCCCTTTGGAAAAGTTGCAAGAAATAGA 616
589 TAAGTTGAGTTGGTGAAGCTTAACTCAGCTATTGCAATGCTTTACCAAGGATGGG 648
617 GCGCGCGAGTACATCTCTTTCTATCAACAAAGATTTCTATPACAAAGACTTTTACTTAA 676
649 AGCTAGAAAATACATATCCATTTACGAAAACATGATGACACACCACTTTGCTTTTGA 708
677 ACTTGTAAATGATGATTTCAATTTGCTTCACTTCACTTCACTTCACTTCACTTCACT 736
709 ATTTGCTAAATGGATTTTAACTGCTGCAAAAGTTTCACTTCACTTCACTTCACTTCACT 768
737 GTGCAATCGTGAAGCTTTCATATCAAGAAAGCGCACTTGTTTAAGATAGAAAT 796
769 TACAAGTTGGTGAAGATTTGGAATTTGCAATAAATATCCATATGCAAGAGCAGGTT 828
797 TGTGAATGCTACTTTTGGGAGTAGGTTAGGCTATGAGCCACAGATTTCCCGGGCTAG 856
829 GGTGTAGTGTACTTCTGATATTTAGAGTGTATTTGAGCCAAATATAGTGTGCGAG 888
857 AGTTTCTTCAAAAAGCTTGTGCTTATTAATCTTATAGATGACACTTATGATGGTA 916
889 AAAAAATGATGACAAAAGTACTCAACCTGACCTCCATTTATGAGACACTTTTGTATGCTTA 948
917 TGTACTTATGAAGAACTTAAAGATCTTACTGAAGCTTTGAAAGGTGGTCAATTTACATG 976
949 TGAACCTTTGACAACTTTGTGACTTTCAATGATGCAATCCAGAGATGGGATGCTAATGC 1008
977 GTTAGACACACTTCCAGAAATACATGAACCGATATACAAATTTATGATGATACATAC 1036
1009 AATTGATTCAATACAACTATATGAGACCTGCTTATCAAGCTTCTTAGACACTTTTACAG 1068
1037 AGAAATGGAAGATTTCTTCAAGGAGGGAAGAACAGATCTATTATTAACCTGGCGCAAGA 1096
1069 TGAATGGAACAGTGTGTCGAAGAGGTAACTGAGCCGCTGATATCTATATGCAAAAA 1128
1097 ATTTGTGAAGAGCTTTGTTAGAAACCTGTGTTGAAGCAAAATGGGCAAAATGA---GGG 1153
1129 TGAGATGAAAAGTTGGTGAAGGCTTATTTAAGAAACCAATGGTTGATGATTGGA 1188
1154 ACACATACCACCACTGAAGAGCATGATCCAGTTGTATATCATTAATCTGGCGGTGTAACT 1213
1189 CCATATTCCAAAATATGAGCAACAAAGTGAAGATGCAATCGTAAGTGTGCTATATGAT 1248
1214 GCTTACAACTGTTTATCTTGGCTAGTGTATATTACAAAGAGTCTGTGCAATG 1273
1249 GATATCAACAACTTGTGTCGATATGAGAAATTTATATCCACGAGACTTTTGAATG 1308
1274 GGCTGTCTGCGACCTCTCTTTTATAGTACTCAGGTATATCTTGTGTCGACGCTTAAATGA 1333

1309 GTTGTGATGAATGAGTCTGTGATTGTTTCGAGCTTCCGCAATTGATTCGAGAGCAATGAACGA 1368
1334 TCTCATGACCCACCAAGCCGAGCAAGAAAGAAACATAGTTTCATCGAGCCTTGAAGTTA 1393
1369 TATTTGTGGACATGAAGATGAACCAAGAAAGAGGACATAGCTTCACTTATTGAATGTTA 1428
1394 TATGAAGAAATATAATCTCAATGAGGAGTATGCCCAACCTTGTATTACAAAGAAAGTAGA 1453
1429 CATGAAGATTTAGGAGCTTCAAGCAAGAGACITACATTAAGTTTCTGAAAGAGGTCAC 1488
1454 AGATGTGTGAAGATATAAACCGAGAGTACTCTCAACACTTAAACATTTCCAAAGCCGTT 1513
1489 CAATGCTATGGAAGGACATAAACAACAATCTCCCGTCCAACTGAAGTACCAATGTTGT 1548
1514 ATTGATGGCTGTGATCTTATTGTCAGTTCCTTGAAGTTCAATATGCAAGAAAGGATAA 1573
1549 CCTTGAAGAGTTCTAAATTTGACACGTGTGGTGACACGTTATATAGGAGAAAGATAC 1608
1574 CTTTCAACGCTATGGGAGACGAATACAAACATCTCATAAAGTCTCTTACTCGTTTATCTCTAT 1633
1609 ATATTCAACGCCCAAGGAAACTTTAAACAATGATTAATCAATACTAATTTGAATCTGT 1668
1634 GAGTATATGAGGAT 1647
1669 CAAAATATAAATAT 1682

RESULT 4
US-09-903-012-51
Sequence 51, Application US/09903012
Patent No. US20020094557A1
GENERAL INFORMATION:
APPLICANT: Chappell, Joseph
APPLICANT: No. US20020094557A1, Joseph P.
APPLICANT: Starks, Courtney M.
APPLICANT: Manna, Kathleen R.
TITLE OF INVENTION: SYNTHASES
FILE REFERENCE: 07678-025001
CURRENT APPLICATION NUMBER: US/09/903,012
CURRENT FILING DATE: 2001-07-11
PRIOR APPLICATION NUMBER: 09/398,395
PRIOR FILING DATE: 1999-09-17
PRIOR APPLICATION NUMBER: 60/100,993
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/130,628
PRIOR FILING DATE: 1999-04-22
PRIOR APPLICATION NUMBER: 60/150,262
NUMBER OF SEQ ID NOS: 58
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 51
LENGTH: 2024
TYPE: DNA
ORGANISM: Lycopersicon esculentum
FEATURE:
NAME/KEY: CDS
LOCATION: (32)...(1675)
OTHER INFORMATION: VFNT germacrene C synthase
US-09-903-012-51

Query Match 21.7%; Score 357.6; DB 9; Length 2024;
Best Local Similarity 52.7%; Pred. No. 2.1e-85;
Matches 861; Conservative 0; Mismatches 749; Indels 24; Gaps 3;
QY 29 TCGCCCCATTGCGCAACTTCTCCAGCATTTGGGAGATCAGTTTCTCATCTATCAAAA 88
DB 58 TCGCCCCTTGGCTAATTTTCCACCCATCTGTTTGGGATATCATTTCTTTTATACTCA 117
QY 89 GCAAGTAGAGCAAGGGTGGAGACAGATAGTGAATTTAAATAAGAGTGGCGCAACT 148
DB 118 TGAATTTACTAATCAAGAAAAGTTGAGTTGATGATCAAGAGACATTTAGAAAAT 177

149 ACTAAGAGAGCTTTGGATATTCCTATGAACATGCCAATTTGTTGAAGCTGATTGATGA 208
178 GCTGGTGAAGACTTTCGCAATATGACATCAAAA-----GCTGTGTTGATGAGCG 228
209 AATTCAAGCCCTTGGAAATACCGTATCACTTTGAAAGGAGATTGATCAATGCAATGCAATG 268
229 GATCAAGATTTGGGAGTGGCTTATCACTTTGATTAATGAAATGAAATGCAATCAATCAAAA 288
269 TATTATGAACAT-----ATGGTGAATCTGGAATGGTGAACCGCTCTTCCTT 316
289 CATTTTGTGATCGATCGTCAACACAGATGATTAATGACAAACACCTTTACGTTGTGCTCT 348
317 ATGGTTCGCTTATCGGAAGCAGGATATTTATGTTACATGATGTTTCAATPACTA 376
349 TCGTTTTCGACTTGTGAGCAACAGGCCCATTTACATGCTTCAGATGTTTCAAGCAAT 408
377 TAAAGACAAAATGGAGCGTTTCAAGCAATCGTTAGCTAATGATGTTGAAGGTTTGGTTGA 436
409 CACCAACCAAGATGGAAATTTCAAGGAACACTTACTAATGATGTTCCAGGATTTAGAG 468
437 GTTGTAGCAAGCACTTCTATGAGGGTACCTGGGAGATTAATTAGAAAGATGCTCTTGG 496
469 TTTGTATGAAGCATCACATCTGAGAGTGGTAATGAGGAGATTTCTTGAAGAGCTCTTAC 528
497 TTTTACAGCATCTCGTCTTAGCATTTATGACAAAGATGCTTTTCTCAAAACCCGCTCT 556
529 ATTTACCAACCATCTCTCGAGTCTATTGTTCTCCAACTTGAACAATAATAAATCTCTCT 588
557 TTTTACCGAAATACAAACGGGCACTAAAGCAACCCCTTTGGAAAAGTTGCCAAGATAGA 616
589 TAAGTTGAAGTTGGTGAAGCTTTAATCAGCTTATTCGATGACTTTTACCAAGGATGG 648
617 GCGGGGAGTACATCTCTTCTATCAACAAAGATCTCTATCAACAGACTTTTACTTAA 676
649 AGCTAGAAAATACATATCAATTTACGAAACCAATGATGACACACCACTTCTTTTGA 708
677 ACTTGCTAAGTTAGTTCAATTTGCTTCACTTTCAGTTCATGACAGGAGAGCTCAGCCAT 736
709 ATTTGCTAATTTGATTTTAAATGCTGCTCAAAAGTTTCCACCAAGAGAGCTTAGTGAT 768
737 GTGCAATGTTGGAAAGCTTTTGCATATCAAGAAAGCAGCTTCTTTAAGAGATAGAT 796
769 TACAAGTTGGTGAAGATTTTGGATTTTGCATTAATTAATCAATATGCAAGAGACAGTT 828
797 TGTGAATGCTACTTTTGGGAGTACTAGTTTCAAGCTATGAGCACAGTATTCCTGGGCTAG 856
829 GGTGAGTGTACTTCTGGATATTAGGAGTGTATTTGAGCCAAAATATAGTCTGCGAG 888
857 AGTTTCTTCAAAAGCTGTTGCTGTTTATTAATCTTATAGATGACACTTATGATGCTTA 916
889 AAAAAATGATGACAAAAGTACTCAACCTGACCTCCATTTATGACGACACTTTTGTGCTTA 948
917 TGGTACTTATGAAGAACTTAAAGATCTTTTACTGAAGCTGTTGAAAGTGGTCAATTACATG 976
949 TGCAACCTTTGACGAACTTTGACTTTCAATGATGCAATCCAGATGGGATGCTAATGC 1008
977 CTTAGACACACTTCCAGAAATACGAAACCGATATACAAATTTATTCGATATACATACAC 1036
1009 AATTGATTTCAATCAACCATATATGAGACCTGCTTATCAAGCTCTTACGATTTTACAG 1068
1037 AGAATGGAAAGATTTCTTGAAGAGGAGGAGAAACAGATCTATTATTAATCTGCGGCAAGA 1096
1069 TGAATGGAAAGATTTGTTGCTTCAAGAGAGTAAACTGGACCGTGTATCTATGCAAAAAA 1128
1097 ATTTGTGAAGAGTTTGTGAGAAACCTGATGTTGAAGCAAAATGGGCAATGA---GGG 1153
1129 TGAGATGAAGAGTTTGTGAGAGCTTATTTAAGGAACCCCAATGTTGCAATGATTTGGA 1188
1154 ACACATACCAACCACTGAAGAGCATGATCCAGTTGTTAATCAATTAATCTGCGGCTGAACCT 1213
1189 CCAATTTCCAAAATATAGGAACCAAGTGGAGATGCAATCGTAAGTCTGGCTATATGAT 1248
1214 GCTTACAAACCACTTGTCTTCTTGGCATGAGTATATTTTCAAAAGAGTCTGTGCAATG 1273

1249 GATATCAAAACATTTGCTTGGTTCGGTATAGAAAGATTTATATCCACGAGACTTTTGAATG 1308
1274 GGCTGTCTCTGACACCTCTCTTTTATAGATATCTAGGTATATCTTGGTCGACGCTAAATGA 1333
1309 GTTGATGAATGAGTCTGTGATTTGTCGAGCTTCCGCAATGATTCGAGAGCAATGAACGA 1368
1334 TCTCATGACCCACAGCCCGAGCAGAAAGAAACATAGTTTCATCGAGCCTTGAAGTTTA 1393
1369 TATTGTTGGACATGAAGATGAACAAAGAAAGAGGACATGTAGCTTCACTTATTGAATGTTA 1428
1394 TATGAAGGAATATAATGTCAATGAGGAGTATGCCCAACCTTGATTTTCAAGGAAGTAGA 1453
1429 CATGAAGATTTAGGAGCTTCAAGCAAGAGACTTACATTAAGTTCTGGAAGAGGTAC 1488
1454 AGATGTGTGGAAGATATAAACCAGAGTACCTCACAACTTAAACCATTTCCAAAGCCGTT 1513
1489 CAATGATGGAAGGACATAAACAACCAATTTCTCCGCTCCAACTGAAGTACCAATGTTGT 1548
1514 ATTGATGGCTGTGATCTATTGTCGAGTTTCTTGAAGTTCAATATGCAAGGAAGATTA 1573
1549 CCTTGAACGAGTTCTAAATTTGACACGCTGGCTGACACGTTATATAAGGAGAAAGATAC 1608
1574 CTTTACACGTTATGGGAGACGAATACAAACATCTCATAAAGTCTCTACTCGTTTATCCTAT 1633
1609 ATATTCAACCCCAAGGAAACTTTAAACATGATTAATCAATACTAATTGAATCTGT 1668
1634 GAGTATATGAGGAT 1647
1669 CAAAATATAATAT 1682

RESULT 5
US-09-900-797-51
; Sequence 51, Application US/09900797
; Publication No. US20030087406A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20030087406A1, Joseph P.
; APPLICANT: Starke, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/900,797
; PRIOR FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 51
; LENGTH: 2024
; TYPE: DNA
; ORGANISM: Lycopersicon esculentum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (32)...(1675)
; OTHER INFORMATION: VFNT germaecene C synthase
US-09-900-797-51

Query Match 21.7%; Score 357.6; DB 10; Length 2024;
Best Local Similarity 52.7%; Pred. No. 2.1e-85;
Matches 861; Conservative 0; Mismatches 749; Indels 24; Gaps 3;
29 TCGCCCATTTGCCAATTTCTCCTCAAGCATTTGGGAGATCAGTTTCTCATCTATCAAAA 88
58 TCGCCCATTTGGCTTAATTTTCAACCATCTGTTTGGGATATCATTTCTTCTTACTCA 117
89 GCAAGTAGAGCAAGGGTGGACAGATAGTGAATTTAAAAAAGAGTGGCGCAACT 148

Db 118 TGAATTACTAATCAAGAAAAAGTTGAATTGATGAGTACAAAGAGACAATTTAGAAAAAT 177
Qy 149 ACTAAAAAGAGCTTTGGATATTTCTTATGAAACATGCCAAATTTGTTGAAGCTGATTGATGA 208
Db 178 GCTGGTGAACAACTTGGACAAATAGCACTCAAAA-----GCTTGTGTTGATAGACGC 228
Qy 209 AATTCAAGCCCTTGGAAATACCGTATACCTTTGAACGGGAGATTGATCATGATTCGAATG 268
Db 229 GATCAACAGATTGGGAGTGGCTTATCATTTCCGATTAATGAAATTTGAAACATCCATTTCAAAA 288
Qy 269 TATTATTGAACAT-----ATGGTGATAACTGGAATGTGTGACCGCTCTTCCTT 316
Db 289 CATTTTGTATGATCGTCCAAACAGAAATGATATGACAAACCTTTAGTGTGTCTCT 348
Qy 317 ATGTTCCGCTTATGCGAAAGCAAGGATATTTATGTTACATGTGATGTTTCAATAACTA 376
Db 349 TCGTTTTCGACTTGTGAGGCAACAGGCCAATTAATGATGCTTCAGATGCTTCAAGCAATT 408
Qy 377 TAAAGACAAAATAGGAGCGTTCAAGCAATCGTTAGCTAATGATGTTGAAGTTTGGTTGA 436
Db 409 CACCAACCAAGATGGGAAATTCAGGAAACACTTAATAATGATGTCCTCAAGGATTTAG 468
Qy 437 GTTGTACGAACCACTTCTATGAGGTTACCTGGGAGATTATATTAGAGATGCTCTTGG 496
Db 469 TTTGTATGAAGCATCACAATCTGAGAGTGGTAAATGAGGAGATCTTGAAGAAAGCTTTAC 528
Qy 497 TTTTACAGATCTCGTCTTAGCATTTATGACAAAAGATGCTTTTCTACAAAACCGCTCT 556
Db 529 ATTTACCACTCATCTCGAGTCTTGTCTCCAACTTGGCAATATAATACTCTCT 588
Qy 557 TTTTACGAAATACACGGGCATTAAGCAACCCCTTTGGAAAAAGTTGCCAAGAAATAGA 616
Db 589 TAAGGTTGAAGTTGGTGAAGCTTAACCTCAGCTATTCGATGACTTTACCAAGGATGG 648
Qy 617 GCGCGCGAGTACATCTCTTCTATCAACAAGATTTCTCATACAAGACTTTACTTAA 676
Db 649 AGCTAGAAAATACATATCAATTTACGAAAACATGATGACACACCACCTTCTTTGAA 708
Qy 677 ACTTGCTAAGTTAGAGTTCAATTTCTTCAGTTCATTGCAACAGGAGAGCTCAGCCATCT 736
Db 709 ATTTGCTAAATTTGATTTTAACTGCTGCAAAAGTTTCAACAAAGAGACTTTAGTATCT 768
Qy 737 GTGCAATGGTGAAGCTTTGATATCAAGAAAGCGCACTGTTTGAAGAGATAGAAAT 796
Db 769 TACAAGTGGTGAAGATTTGGATTTTGCATAATTAATATCCATATGCAAGAGACAGGTT 828
Qy 797 TGTGAATGCTACTTTTGGGAGTAGTTTCAAGCTATGAGCCACAGTATTCCTCGGCTAG 856
Db 829 GGTGAGTGTACTTCTGATATTAGAGTGTATTTTGGAGCCAAATATAGTCGTGCGAG 888
Qy 857 AGTTTCTTCAAAAAGCTTTGCTGTTTATAACTCTTTATAGATGACACTTATGATCGTA 916
Db 889 AAAAATGATGACAAAAGTACTCAACTGACCTCCATATTATGACGACACTTTTGTGCTTA 948
Qy 917 TGGTACTTATGAAGAACTTAAATCTTTTACTGAAGCTGTTGAAGTGTGTCATTTACATG 976
Db 949 TGCACCTTTGACGAATTTGTGACTTTTCAATGATGCAATCCAGAGATGGGATGCTAATGC 1008
Qy 977 CTAGACACACTCCAGATACATGAACCGATATACAAATATTCTCATGATACATACAC 1036
Db 1009 AATTGATTCATACAAACATATAGACCTGCTTATCAAGCTCTTCTAGACATTTTACAG 1068
Qy 1037 AGAAATGGAAGATTTCTTGAAGAGGAGGAAGAACAGATCTATTATTAAGTGGCGCAAGA 1096
Db 1069 TGAATGGAACAAAGTTGTTCCAAAGAAAGGTAACCTGGACCGTGTATATATGCAAAAAA 1128
Qy 1097 ATTTGTGAAGATTTGTTAGAAAACCTGATGTTGAGCAAAATGAGGCAAAAGA----GGG 1153
Db 1129 TCAGATGAAAAGATTTGGTGAGAGCTTATTTTAAAGGAAACCCCAATGTTGAATGTTGA 1188
Qy 1154 ACACATACCAACCACTGAGAGCATGATCCAGTTGTAATCATTTACCTGCGGCTTAACCT 1213
Db 1189 CCATATTCAAAATATGAGAAACAAGTGGAGAAATGCAATCGTAAGTGTCTGTATATGAT 1248

Qy 1214 GCTTACAAACAACCTTGTATTTCTTGGCATGAGTATATATTCACAAAAGAGTCTGTGCAATG 1273
Db 1249 GATATCAACAACCTTGTCTTGGTCGGTATAGAAGAAATTTATATCCACGAGACTTTTGAATG 1308
Qy 1274 GCGTGTCTCGACCTCTCTTTTATAGTACTCAGGTATACTTGGTTCGACGCTTAATGA 1333
Db 1309 GTTGATGAATGAGTCTGTGATTTGAGCTTCGGCAATGATTCGCCAGAGCAATGAACGA 1368
Qy 1334 TCTCATGACCCCAAGGCGGAGCAAGAAAGAAACATAGTTTCATCGAGCCTTTGAAAGTTA 1393
Db 1369 TATTGTTGGACATGAAGATGAACAAGAAAGAGACATGTAGCTTCACTTATTGAATGTTA 1428
Qy 1394 TATGAAGGAATATAATGTCAATGAGGAGTATGCCAAACCTTTGATTTTACAAGAGATAGA 1453
Db 1429 CATGAAAGATTAATGGAGCTTCAAGCAAGAGACTTACATTAAGTTCCTGAAAGAGGTCA 1488
Qy 1454 AGATGTGCGAAGATATAAACCAGAGTACCTCAACAATTAACAAATTCACCAAGCCCTT 1513
Db 1489 CAATGATGAAGGACATAAACAACAATTTCTCCGTCCACTGAAGTACCATGTTTGT 1548
Qy 1514 ATTGATGGCTGTGATCTATTGTCGCAAGTTTCTTGAAGTTCAATATGCAAGAAAGATAA 1573
Db 1549 CCTTGAACGAGTTCTAAATTTGACACGTTGGCTGACACGTTATATAAGGAGAAAGATAC 1608
Qy 1574 CTTCACACGTATGGGAGACGAATACAAACATCTCATAAAGTCTCTACTCGTTTATCCTAT 1633
Db 1609 ATATTCAACGCCCAAGGAAACTTAAAAACATGATTAATCCATACTAATTTGAATCTGT 1668
Qy 1634 GAGTATATGAGAT 1647
Db 1669 CAAAATATAAATAT 1682

RESULT 6
US-09-893-820-51
; Sequence 51, Application US/09893820
; Publication No. US20040053386A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20040053386A1, Joseph P.
; APPLICANT: Sparks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/893,820
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US/09/398,395A
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 51
; LENGTH: 2024
; TYPE: DNA
; ORGANISM: Lycopersicon esculentum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (32)...(1675)
; OTHER INFORMATION: vFNT germacrene C synthase
US-09-893-820-51

Query Match 21.7%; Score 357.6; DB 13; Length 2024;
Best Local Similarity 52.7%; Pred. No. 2.1e-85;
Matches 861; Conservative 0; Mismatches 749; Indels 24; Gaps 3;
Qy 29 TCGCCCCATTCGCAACTTCTCCCAAGCATTTGGGAGATCAGTTTCTCATCTATCAAAA 88

Db 58 TCGCCCTGGCTAAATTTTCCACCATCTGTTTGGGATATCATTTCTCTTTTATACTCA 117
Qy 89 GCAAGTAGAGCAAGGGTGGAAACAGATAGTGAATGATTTTAAAAAAGAGAGTGGGCAACT 148
Db 118 TGAATTTACTAATCAAGAAAAGTTGAAGTTGATGAGTACAAAGAGACAAATTAGAAAAT 177
Qy 149 ACTAAAGAACCTTTGGATATCTCTATGAACATGCCAATTTGTTGAAGCTGATGATGA 208
Db 178 GCTGGTGGAAACTTTGCGCAATAGCACTCAAAA-----GCTTGTGTTGATAGACGC 228
Qy 209 AATTCAACGCCCTTGAATACCGTATCACTTTGAACGGGAGATTGATCATGCAATTCGAATG 268
Db 229 GATCGAAGATTGGAGTGGCTTATCATTTGATATGAAATGAAACATCCATTCAAAA 288
Qy 269 TATTTATGAACAT-----ATGGTGAATGAACTGGAATGGTGAACCGCTCTTCTCT 316
Db 289 CATTTTGTATCATCGTCCAAACAGATGATATGACAAACACCTTTTACGTTGTGCTCT 348
Qy 317 ATGTTTCGCTTATGCGAAAGCAGGATATGTTTATCATGTGATGTTTCAATAACTA 376
Db 349 TCGTTTTCGACTTGTGAGGCAACAGGCCATTTACATGTCTTCAGATGTGTTTCAAGCAAT 408
Qy 377 TAAAGACAAAATGGAGGTTTCAAGCAATCGTTAGCTAATGATGTTGAAGTTTGCCTTGA 436
Db 409 CACCAACCAAGATGGGAATTCAGGAAACACTTACTAATGATGTTCCAAAGATTTATTGAG 468
Qy 437 GTTGTACGAACCACTTCTATGAGGTTACCTGGGAGATTTATTTAGAGATGCTCTTGG 496
Db 469 TTTGTATGAACATCACATCTGAGAGTGGTAAATGAGGAGATTTCTTGAAGAGCTTTAC 528
Qy 497 TTTTACAGATCTGCTTTAGATTTATGACAAAAGATGCTTTTCTACAAACCCGCTCT 556
Db 529 ATTTTACCACCACTCATCTCGAGTCTATTGTCTCCAACTTTGACAAATAATAAATCTCT 588
Qy 557 TTTTACCGAATACAAACGGGCACTAAAGCAACCCCTTTGGAAAAGTTGCCAAGATAGA 616
Db 589 TAAGTTGAAGTTGGTGAAGCCTTAACCTGAGCTATTGCGATGCTTTTACCAGAGATGG 648
Qy 617 GCGCGCGAGTACATTTCTTCTATCAACAACAGATTTCTATACAACAGACTTTACTTAA 676
Db 649 AGCTAGAAAATACATATCCATTACGAAACCAATGATGACACACCACTTCTTTTGA 708
Qy 677 ACTTGCTAAGTAGATTTCAATTTGCTTCACTGATGCAACAGGAGAGCTCAGCATGT 736
Db 709 ATTTGCTAAAATGGATTTTAAATGCTGCTGCAAAAGTTTCAACCAAGAGAGCTTAGTATCT 768
Qy 737 GTGCAATGTTGGAAAGCTTTTTCGATATCAAGAGAACGACCTTTTAAAGATAGAT 796
Db 769 TACAAGTTGGTGAAGATTTGGATTTTGGCAATTAATATCCATATGCAAGAGACAGGT 828
Qy 797 TGTGAATGCTTACTTTTGGGACTAGGTTCAGGCTATGAGCCACACAGTATTTCCGGGCTAG 856
Db 829 GGTGAGTGTACTTCTGATATTAGGAGTGTATTTTGGAGCCAAATATAGTCTGTCGAG 888
Qy 857 AGTTTCTTACAAAGCTGTTGCTGTTATTAATCTTTATAGATGACATTTATGATGGTA 916
Db 889 AAAAAATGATGACAAAGTACTCAACCTGACCTCCATTTATGACGACACTTTTGTATGCTTA 948
Qy 917 TGGTACTTTATGAACACTTATGATCTTTTACTGAAGCTGTTGAAAGTGGTCAATTACATG 976
Db 949 TGCACCTTTGACGAACTTTGACTTTCAATGATGCAATTCAGAGATGGGATGCTAATGC 1008
Qy 977 CTTAGACACACTTCCAGAAATACATGAAACCGATATACAAATTTATTCATGATACATACAC 1036
Db 1009 AATTGATTTCAATACACCATATATGAGACCTGCTTTATCAGCTCTTCTAGACATTTACAG 1068
Qy 1037 AGAAATGAAGAAATTTCTTGAAGAGGAGGAGAAACAGATCTATTATTAATCGCGCAAGA 1096
Db 1069 TGAATGGAACAAGTGTGTGCAAAAGAGGTAAACTGTGACCGCTGATATCTATGCAAAAAA 1128
Qy 1097 ATTTGTGAAGAGTTTGTAGAACTGATGTTGAAAGCAAAATGGGCAAAATGA---GGG 1153
Db 1129 TGAGATGAAAAGTTGGTGAGAGCCTATTTTAAGGAAACCCCAATGGTTGAATGATTTGTA 1188

Qy 1154 ACATATACCAACCACTGAGAGCATGATCCAGTTGTATCTTATCTTACTTGGCGGTGCTAACCT 1213
Db 1189 CCATATTCCAAATATGAGGACCAAGTGGAGAAATGCAATCGTAAAGTCTGGCTATATGAT 1248
Qy 1214 GCTTACAAACACTTTGTTATCTTGGCATGAGTGAATATTTTCAAAAAGAGTCTGTGCAATG 1273
Db 1249 GATATCAACAACTTGTCTGCTGCGTATAGAAGAAATTTATATCCACGAGACTTTTGAATG 1308
Qy 1274 GGCTGTCTGTCACCTCTCTTTTATAGATCTCAGGTATATCTTGGTCGACGCTAAATGA 1333
Db 1309 GTTGATGAATGAGTCTGTGATTTGTTCCGCTTCCGATTTGTTCCAGAGCAATGAACA 1368
Qy 1334 TCTCATGACCCACAAGCCGAGCAAGAAAGAAAACATAGTTTCATCGAGCCTTGAAGTTA 1393
Db 1369 TATTTGTGACATGAGATGAACAAGAAAGAGGACATGTAGCTTCACTTATTTGAATGTTA 1428
Qy 1394 TATGAAGATATTAATGTCAATGAGGAGTATGCCAAACCTTGTATTTACAAGGAAGTAGA 1453
Db 1429 CATGAAAGATTTATGGAGCTTCAAAAGCAAGAGACTTACATTTAAGTTCTTGAAGAGGT 1488
Qy 1454 AGATGTGTGAAAGATATAAACCAGAGAGTACCTCACAACATAAAACATTTCCAAGGCCGTT 1513
Db 1489 CAATGATGAGAGACATAAAACAAACATTTCTCCGTCCTCAACTGAAGTACCAATGTTGT 1548
Qy 1514 ATTTGATGGTGTGATCTATTTTGTGCCAGTTTCTTGAAGTTCAATATGCGAGAAAGATAA 1573
Db 1549 CCTTGAACGAGTTCTAAATTTTGACACGTGTGGCTGACACGTTATATAGGAGAAAGATAC 1608
Qy 1574 CTTCACAGTATGGGAGACGAATAACAACATCTCATAAAGTCTCTACTCGTTTATCTTAT 1633
Db 1609 ATATTCAACCGCCAAAGGAAACCTTTAAAACATGATTTAATCCAAATACTAATTTGAATCTGT 1668
Qy 1634 GAGTATATGAGGAT 1647
Db 1669 CAAAATATAATAT 1682

RESULT 7
US-10-424-599-43432/c
; Sequence 43432, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 43432
; LENGTH: 1937
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_139218C.1
US-10-424-599-43432

Query Match 21.3%; Score 351.6; DB 13; Length 1937;
Best Local Similarity 53.0%; Pred. No. 8.6e-84;
Matches 865; Conservative 0; Mismatches 734; Indels 33; Gaps 4;
Qy 27 ATTCGCCCACTTCCCAACTTTCTTCCAAAGCAATTTGGGAGATCAGTTTCTCATCTATCAA 86
Db 1815 ATTCGGAACACTGCAAAATTTTCTCTTCCGTTTGGGCGGATTTATTTCTTTACTATGTT 1756
Qy 87 AAGCAAGTAGACCAAGGGTGGACAGATGATGATTTTAAAAAAGAGAGTGGCGCA 146
Db 1755 CCTAGTTCCGTGGAAGATGATGATCATCAACCAAGCTCAATCAACAAAAGAGAGTGTG 1696

QY 147 CTACTAAAGAGCTTTGGATATCTCTATGAACATGCCCAATTTGTTGAAGCTGATGAT 206
Db 1695 AGGAAGATGCTTATTGCTCTTATGATAACAATTTCTATTTTAAATTAGAATTTATTGAC 1636
QY 207 GAAATTCACCGCTTTGGAATACCGTATCACTTTGAAACGGGAGATTGATCATGCAATTCGCA 266
Db 1635 TCAGTCCAAACGCTTGGGTGTTCTTACCACTTTTGAACATGAAATTCATGGAGCAATTCAC 1576
QY 267 TGTATTTATGAACAATATGGTGATTAACCTGGAATGG-----TGAC 305
Db 1575 CAAATTTTCAACAATTTCAACAAGGACCAATAATATCATAACTCACGATGATGATCTTTGT 1516
QY 306 CGCTCTTCTTATGTTGCTGCTTATGCGAAACGACGAGGATATATGTTTACATGATGATGTT 365
Db 1515 CATGGGCTTTACTCTTTTGGTGTCTTAGGCAACAGGTTATCACTTTTCAATCAAAATGTA 1456
QY 366 TTCAATAACTATAAAGACAAAATGGAGCGTTCAAGCAATCGTTAGCTAATGATGTTTGA 425
Db 1455 TTTTCAAAATTCAAAGACCAAACTAGAACTTTAGTGAAGAAAGCAGCCCAATGATATACAA 1396
QY 426 GGTGTGCTTGTAGTTGACGAACCACTTCTATGAGGTTACCTGGGGAGATATATTTAGAA 485
Db 1395 GGAATGTTGAGCTTTGACGAAGCTGCTGAACCTAAGAATGATGAGAGGATATACTTGAA 1336
QY 486 GATGCTCTTGGTTTACACGATCTCGTCTTAGCAATATGACAAAAGATGCTTTTCTACA 545
Db 1335 GAAGCACAAATTTTGGCTTCTGCTTCAAGTTAACTAAGTCTCTAATACCCAAAT-----A 1282
QY 546 AACCCCGCTCTTTTACCGAATAACAACGGGCACTAAAGCAACCCCTTTGGAAAGGTTG 605
Db 1281 AGTCCTTCTATGATGTCACAAGCTCAAGCATAGCTTTAAGACGATCACTTGGAAAGGATG 1222
QY 606 CCAAGATAGAGCGGCGGAGTACATTCCTTTCTATFCAACAACAAGATTCCTCATACAAG 665
Db 1221 CCTAGGCTGGAGGCCACATATATATATGCTTTCTACGAAGAAGATTCCTTACATGATGAA 1162
QY 666 ACTTTACTTAACTGCTAAGTTAGATTTCAATTTGCTTCAGTCAATGACACAGGAGAG 725
Db 1161 AAATGCTTAACTTTGCAAACTAGATTTTAAATGTTGCGAGAGTTTACATCAGAAGAA 1102
QY 726 CTCAGCCATGTTGCAAAATGGTGG---AAAGCTTTTCGATATCAAGAAGAACGACCTTGT 782
Db 1101 GTTAAACATGTGACCAAGTGGTGGATTAAGATTTTAAATGTTCTCAACGAATTTGCCATTT 1042
QY 783 TTAAGATAGAAATTTGTAATGCTATTTTGGGACCTAGGTTTACGCTATGAGCCACAG 842
Db 1041 GTACGATAGGATGCAAGATGTTTACTTTTGGATTTTGGGAATATACCTTTGAGCCACAA 982
QY 843 TATTCCCGGCTAGAGTTTCTTCAAAAAGCTGTTGCTGTTATTAATCTTTATAGATGAC 902
Db 981 TATTCTCTTGTAGAAAGGATAACAACGAAAGTAATTTGCTCTATGTTCTGTCTATGATGAC 922
QY 903 ACTTATGATGGTATGTTACTTATGAAGAACTTAAAGATCTTTTACTGAAAGCTGTTGAAAG 962
Db 921 ATGTATGATGCTATGGAACCACTTGACGAACCTTGAGCTTTTCCCAATGCAATGAGAGG 862
QY 963 TGGTCAATTAATGCTTTAGACACACTTCCAGAAATACATGAACCGGATATACAAATTTATTC 1022
Db 861 TGGGATATTTGTTGCTGGATGATCTCCAGAAATACATGAAAGTATGTTATATAGAAAT 802
QY 1023 ATGGATACATACAGAAATGAAGAAATTTTTCGAAAGGAGGAGAACAGATCTATTT 1082
Db 801 TTGAATGTTTATGAAGAATAGAGGAAGATGAGAAAACAAGGAAAGTATATTGTCATC 742
QY 1083 AACTGCGCAAGAAATTTGTGAAGAGTTTGTAGAAAACCTGATGGTTGAAGCAAAATGG 1142
Db 741 AAGTATGCTAAGAAAGATGAAGATTAATCAAGGCTCACATGGCTGAGGCAAGATGG 682
QY 1143 GCAATAGGGGACACATACCAACCACTGAAGAGATGATCCAGTTGTATATCAATCTGCG 1202
Db 681 CTTCAATTCATACACCGCTCAATAGAGGAGTACATGCAAGTTAGAAATGATCAAGT 622
QY 1203 GGTGCTAACCTGCTTACACAACTTTGTTATCTTGGCATGATGATATATTCACAAAGAG 1262

Db 621 GGTTACTCTATGGTGATCACCATATGTTTGTGGCATGAAGATACAC---ACAGAGGAG 565
QY 1263 TCTGTGAAATGGGCTGTCTCTGACACCTCCTCTTTTATAGATCTCAGGTATATCTTGTGCGA 1322
Db 564 GTCCTTATATGGGCAACAAGTGATCCCAATAATTTATGGGCTGCTTCAATTTATTTAGG 505
QY 1323 CGCCTAAATGATCTCATGACCCCAAGCCGAGCAAGAAAGAAACATAGTTCTATCGAGC 1382
Db 504 CTTATGATGACATTTGTTGGAATAGTTTGAGCAGGAAAGACACGTTGTCATCAAGC 445
QY 1383 CTTGAAAGTTTATATGAAGGAATAATGTCATGAGGAGTATGCCCAACCTTTGATTTAC 1442
Db 444 ATTGAAGCTTATATGAAGCAACATACACCTCAAGGCAAGATGCCATTAATAAATCTACT 385
QY 1443 AAGGAAGTAGAAGATGTTGGAAGATATATAACCGAGAGTACCTCACAACTAAAAACATT 1502
Db 384 GAGATGTTTAAAGATGCTTGGAGGACATCAATGAGGATGCTTTAATCTTACTGAAGTG 325
QY 1503 CCAAGCCGCTTATGATGGCTGTGATCTATTTTGGCCAGTTTCTTGAAGTTCAATATGCA 1562
Db 324 CCAATGAATTTCTTTTGGCTGTTGTCACCTTGTGCGCATGATGACGTGCTTTTACAAA 265
QY 1563 GGAAGGATTAACCTTACACGATGATGGGAGACGAATACAAAACATCTCATAAAGTCTCTACTC 1622
Db 264 GATGAAGATTAATATACAAATGCAAGGAGGTTAATGAAGATTAATCAAACTTTTATTA 205
QY 1623 GTTTATCTCTATG 1634
Db 204 GTTAAATAAGATG 193

RESULT 8
US-09-887-586A-31
; Sequence 31, Application US/09887586A
; Patent No. US20020094556A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20020094556A1, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/887,586A
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 1944
; TYPE: DNA
; ORGANISM: Solanum tuberosum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (57)...(1724)
; OTHER INFORMATION: vetispiradiene synthase
US-09-887-586A-31

Query Match 20.0%; Score 330.4; DB 9; Length 1944;
Best Local Similarity 52.0%; Pred. No. 4.5e-78;
Matches 857; Conservative 0; Mismatches 761; Indels 30; Gaps 4;

QY 14 AGAAGAAAACCTTATTCGCCCATTCGCAACTTTCCTCCAGCACTTTGGGAGATCAGTT 73
Db 92 AGAGGAGGAGATGTTGCGCCCATAGTACTTCTCTCCAAGTCTTTCGGGTGATCGTTT 151
QY 74 TCTCATCTATCAAAAGCAAGTAGAGCAAGGGTGGAGAGATGATGATTTTAAAAA 133

QY	1199	TGGCGGTGCTAACTCTGCTTACAACTTGTATCTTGGCATGAGTGATATATTCACAAA	1258
Db	1283	CAGCAGCTATTACTTGTCTAGCAGTACATCTTATTTGGGCATGAAGTCTGCTAACAAAGCA	1342
QY	1259	AGAGTCTGTGCAATGGGCTGCTCTGCACTTCTCTTTTGTAGATCTCAGGTACTATTGG	1318
Db	1343	AGATTTTG---AATGGTTGGCAAGAACCTTAAATTTCTTGAGGCTAATGTGACGTTATG	1399
QY	1319	TCGACGCGCTAAATGATCTCATGACCCCAAGCCGAGCAAGAAAGAAACATAGTTTCATC	1378
Db	1400	CCGAGTCATAGATGACATAGCACCTATGAGGTTGAGAAAGGTTAGAGGTGAGATTGCCAC	1459
QY	1379	GAGCCTTGAAGTTATATGAAGGAATATAATGTCATGAGGAGTATGCCCAACACCTTGAT	1438
Db	1460	TGGAATTTGAATTTTACATGAGAGATTATGGTGATATCCACAGAAAGGCCATGGAAAAATT	1519
QY	1439	TTACAAGGAAGTAGAAGATGTTGGAAAGATATATAACCGAGAGTACTCTCAACTATAAAA	1498
Db	1520	CCAGAATAATGGCTGAGACAGCATGGAAGGATGTAATGAAGNAATCTTCGACCACTCC	1579
QY	1499	CATTCCAGGCGGTTATTGATGGCTGTGATCTATTGTGCCAGTTCTTTCGAAGTTCAATA	1558
Db	1580	CGTCTCTACAGAGATTCTCACTCGCATTTCTCAATCTTGTCTCGCATATCGATGTACTTA	1639
QY	1559	TGCAGGAA---AGGATACTTACACGGTATGGGAGACGAATACAAACATCTCTAAAGTTC	1615
Db	1640	TAGACAATCAAGATGATACACTCATCCGGAATAAGTACTAAACCTCATATTATTTC	1699
QY	1616	TCCTACTCGTTTATCCTATGAGTATATGA	1643
Db	1700	GTTCTGGTGGACTCTATTGAAATTTAA	1727

RESULT 9
 US-09-903-012-31
 ; Sequence 31, Application US/09903012
 ; Patent No. US20020094557A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Chappell, Joseph
 ; APPLICANT: No. US20020094557A1, Joseph P.
 ; APPLICANT: Starks, Courtney M.
 ; APPLICANT: Manna, Kathleen R.
 ; TITLE OF INVENTION: SYNTHASES
 ; FILE REFERENCE: 07678-025001
 ; CURRENT APPLICATION NUMBER: US/09/903,012
 ; CURRENT FILING DATE: 2001-07-11
 ; PRIOR APPLICATION NUMBER: 09/398,395
 ; PRIOR FILING DATE: 1999-09-17
 ; PRIOR APPLICATION NUMBER: 60/100,993
 ; PRIOR FILING DATE: 1998-09-18
 ; PRIOR APPLICATION NUMBER: 60/130,628
 ; PRIOR FILING DATE: 1999-04-22
 ; PRIOR APPLICATION NUMBER: 60/150,262
 ; PRIOR FILING DATE: 1999-08-23
 ; NUMBER OF SEQ ID NOS: 58
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 31
 ; LENGTH: 1944
 ; TYPE: DNA
 ; ORGANISM: Solanum tuberosum
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (57)...(1724)
 ; OTHER INFORMATION: vetispiradiene synthase
 US-09-903-012-31

Query Match	20.0%	Score 330.4;	DB 9;	Length 1944;
Best Local Similarity	52.0%	Pred. No. 4.5e-78;		
Matches 857;	Conservative	0;	Mismatches 761;	Indels 30;
Gaps	4;			

QY 14 AGAGAAAAAAGCTATTCGCCCATTTCTCCAAAGCATTTGGGAGATCAAGTT 73

Db 92 AGAGGAGGAGATTGTTGGCCCATAGCTGACTCTCTCCAAAGTCTTTGGGGTGATCGTTT 151
QY 74 TCTCATCTATCAAAAGCAAGTAGACGAGGGTGGACAGATAGTGAATGATTTAAAAA 133
Db 152 TCATTCAATTCCTCCCGCAATCAGATGCTGGAAAATATGCTCAAGAGATCGAAACTTT 211
QY 134 AGAAGTGGCGCAACTACTAAAAAGAGCTTTGGATATTCCTATGAACATGCCCAATTTGTT 193
Db 212 GAAGAAACAATCAAGAAATATATATCTGCACTCTCTCGAAGAACATTTGGCTGAGAAAT 271
QY 194 GAAGCTGATGTAGTAAATTCACCGCTTGGAAATACCGTATACATTTGAAACGGGAGATTGA 253
Db 272 GGATCTGATAGACATTTGTTGAGCGCTTGGCATTTGCTTATCATTTTGAAGAAACAAATAGA 331
QY 254 TCATGCATTCGAATGATTTATTAAGAACATATGTT-----GATAACTGGAA 298
Db 332 TGATATGTTGGATCAATTTTAAAGCAGATCTTAACCTTTGAGGCTCACGAGTACATGA 391
QY 299 TGGTGACCGCTCTTCCTTATGTTTCCGTTCTTATGCGAAAGCAAGGATATTTATGTTACATG 358
Db 392 TTTACAAACTTTATCGTTCAATTTGCACTTTGAGACAAACATGGTTTACAATATCTCCCC 451
QY 359 TGATGTTTCAATACTATAAGACAAATAAGAGCGGTTCAAGCAATCGTTAGCTAATGA 418
Db 452 AAAAATTTTATTTAGATTCGAAGATCAAAAGGCAAAATTTAAAGAAATCTCTTTGTAACGA 511
QY 419 TGTGAAGGTTTGTGCTGTGACGAACAATCTTATGAGGTAACCTGGGGAGATTAT 478
Db 512 CATCAAGGCTCTTTGAACTTATCAAGGCTCGCATGTAAGGACTCATGGAAGAAATAT 571
QY 479 ATTAGAAGATGCTCTTTGGTTTACAGATCTCTGTTTATGCAATTTATGACAAAAGATGCTTT 538
Db 572 TTTGAAGAGGCACATGCTTTCTCTACTGCTCATCTTGAATCT-----GCAGCTCC 622
QY 539 TTTCTACAAACCCGCTCTTTTACCGAAATACAAACGGGCACTAAAGCAACCCCTTTGGAA 598
Db 623 ACATTTGAAGTCACTCTGAGTAAGCAAGTGACACATGCGCCCTTGAGCAATCTCTCCATAA 682
QY 599 RAGGTTGCCAAGATAGAGCGCGCAGTACATCTCTTCTATCAACAAAGATTTCTCA 658
Db 683 GAGCAATCCAGAGTTGAGACAGCGTACTTCTCTATCTAGACAGTATAGAGGAGAACAGAA 742
QY 659 TAACAAGACTTTACTTTAAACTTGTCTAAGTTAGAGTTCAATTTGCTTCAGTCAATGACAA 718
Db 743 GAATGATGTTGCTTCAATTTGCAAACTGGACTTCACTTACTTACTTACAGATGTTGACAA 802
QY 719 GGAAGAGCTCAGCCATGCTGTCMAATGTTGGGAAAGCTTTGATATCAAGAAAGCGACC 778
Db 803 ACAAGAACTTAGTGAAGTATCAAGGTGGTGGAAAGATTTGGATTTTGTGACAAACACTTCC 862
QY 779 TTGTTTAAGAGATAGAAATTTGTAAGTCTACTTTTGGGACTAGGTTTCAGGCTATGAGCC 838
Db 863 ATATGCTAGGATAGAGCAGTGAATGCTACTTTTGGACGATGGGGGTGATGCTGAACC 922
QY 839 ACAGTATTCCTGGGCTAGAGTTTCTTCAAAAAGCTTTGCTGTTTATAACTCTTATAGA 898
Db 923 TCAATACTCTCAGGCTCGTGCATGCTTGTCTAAGACTATAGCAATGATTTCTATAGTAGA 982
QY 899 TGACACTTATGATGCTGATGTTACTTATGAAAGCTTTAAGACTTTTACTGAAAGCTGTTGA 958
Db 983 TGACACATTCGATGCTTATGGCATTTGCAAGAACTTGAGATCTACACCGATGTCATACA 1042
QY 959 AAGTGGTCAATTTACATGCTTAGACACACTTCCAGATATACGAAACCGATATACAAAT 1018
Db 1043 GAGGTGGGATTTAGCCAAATGATCGGCTCCCTGATTAATGAAATCAGTTTACAAAGC 1102
QY 1019 ATTCAATGATATACATACAGAAATGGAAGAAATTTCTTGAAGAGGAGGAAGAAACAGATCT 1078
Db 1103 ACTTTTATGATCTCTCAATGATTTATGAAATGGAGTTGTCCAGAGATGTTAGTCTGATGT 1162
QY 1079 ATTAACTCGGCAAGAAATTTGTGAAGAGTTTGTGTAAGAACTGATGTTGTAAGCAAA 1138
Db 1163 TGTTCCTACTACCGCAAGAAAGAAATGAAAGAAATCGTGAGAAACTATTTTGTGGAAGCAAA 1222

QY 1139 ATGGCAAAATGAGGACACATACCAACCACTGAAGAGCATGATCCAGTTGTTAATCATTTAC 1198
Db 1223 ATGTTTCATTGAAGGATATATGCCCGCAGTCTCTGAGTATCTTAGCAATGCAATTAGCTAC 1282
QY 1199 TGGCGGTGCTAACTGCTTACAACTTGTGTTATCTTGGCATGAGTATATATTACAAA 1258
Db 1283 CAGCACTTATTTACTTGTCTTACGACTCATCTTATTTGGGCATGAAAGTCTCTGTAACAAGCA 1342
QY 1259 AGAGTCTGTCGAATGGGCTGCTCTGCACCTCTCTCTTTTAGATATCTCAGGTATATCTGG 1318
Db 1343 AGATTTTG---AATGGTTGGCAAGAACCTTAAATTTCTTGAGGCTAATGTGACGTTATG 1399
QY 1319 TCGACGCTTAATGATCTCATGACCCCAAGCCGAGCAAGAAAGAAACATAGTTTCATC 1378
Db 1400 CCGAGTCATAGATGACATAGCCACTATGAGGTTGAGAAGGGTAGAGGTCAGATTGCCAC 1459
QY 1379 GAGCTTTGAAAGTTATATGAAGGAATATAATGTCAATGAGGAGTATGCCCAACCTTGAT 1438
Db 1460 TGGAAATGAAATGTTTACATGAGAGATTATGGTGTATCCACGAAAAGGCGCATGGAATAAT 1519
QY 1439 TTACAAGAAAGTAGAAGATGTTGGAAGATATAAACCGAGAGTACCTCACAACTAAAAA 1498
Db 1520 CCAAGAAATGCTGAGACAGCATGGAAGGATGTAATGAAGAAATCCTTCGACCACTCC 1579
QY 1499 CATTCCAAGGCGCTTATTTGATGGCTGTGATCTATTTTGCCAGTTTCTTTGAAGTTCAATA 1558
Db 1580 CGTCTCTACAGAGATTCTCACTCGCAATTTCTCAATCTCTGCTCGCAATATCGATGTTACTTA 1639
QY 1559 TGCAGGAA---AGGATAACTTCACAGTATGGGAGACGAATACAAACATCTCATAAAGTC 1615
Db 1640 TAAGCAAAATCAAGATGATACACTCATCGGAAAAGTACTTAAACCTCATATATTATTCG 1699
QY 1616 TCTACTCGTTTATCTATGAGTATATGA 1643
Db 1700 GTTCTGGTGACTCTATTGAAATTTAA 1727

RESULT 10

US-09-900-797-31
; Sequence 31, Application US/09900797
; Publication No. US20030087406A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20030087406A11, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; FILE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/900,797
; PRIOR FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 1944
; TYPE: DNA
; ORGANISM: Solanum tuberosum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (57)...(1724)
; OTHER INFORMATION: vetispiradiene synthase
US-09-900-797-31

Query Match 20.0%; Score 330.4; DB 10; Length 1944;
Best Local Similarity 52.0%; Pred No. 4.5e-78;
Matches 857; Conservative 0; Mismatches 761; Indels 30; Gaps 4;

QY 14 AGAAGAAAACTATTGCGCCCATTTGCCAACTTTCTCAAGCATTTGGGGAGTCAGTT 73
Db 92 AGAGGAGGAGATTGTTGCGCCCATAGCTGACTTCTCTCCAAGTCTTTGGGGTGAATGTT 151
QY 74 TCTCATCTATCAAAAGCAGTAGAGCAAGGGTGAACAGATAGTGAATGATTTAAAAA 133
Db 152 TCATTTCATTCCTCGCAATCAATTCGCTGGAATATATGCTCAAGAGATCGAACTTT 211
QY 134 AGAAGTCGGCAACTACTATAAAGAGCTTTGGATATTCATGAAACATGCCAAATTTGT 193
Db 212 GAAGGAACAATCAAGAAATATATATCTGCTCTCTCGAAGAACTTTGGCTGAGAAAT 271
QY 194 GAAGCTGATGATGAATTAACAGCTTTGGAATACCGTATATCACTTTGAACGGGAGATGA 253
Db 272 GATCTGATGACATGTTGAGCGCTTTGGCATTCCTTATCATTTGAAAAACAATAGA 331
QY 254 TCATGCAATGCAATGATTTATTAAGAAACATATGCT-----GATAACTGGA 298
Db 332 TGATATGTTGGATCAATTTTAAAGAGAGATCTTAACCTTTGAGGCTACGAGTACAATGA 391
QY 299 TGGTGACCGCTCTCTTATGCTTCGCTTATGCGAAAGCAAGATATATGTTACATG 358
Db 392 TTACAAACTTTATCCGTTCAATTTGCACTATTGAGACAACAATGTTTACATATCTCCC 451
QY 359 TGATGTTTCAATAACTATAAAGCAAAAAATGAGCGTTTCAAGCAATCGTTAGCTAATGA 418
Db 452 AAAAATTTTATAGATTCGAAGTGAAGAAAGGCAAAATTAAGAAATCTCTTTGTAACA 511
QY 419 TGTTGAGGTTGCTTGAGTTGAGAGCAACTCTATGAGGGTACCTGGGGAGATAT 478
Db 512 CATCAAGGCTCTTTGAACTTATACGAAGCTCGCATGTAAGGACTCATGGAAGAGATAT 571
QY 479 ATTAGAAGATCTCTTGTTTATACAGATCTGCTTATGCAATTTAGCAAAAGAGTCTTT 538
Db 572 TTTGGAAGAGCACTTGCTTCTCTGCTGCTCATCTGAACT-----GCACTCC 622
QY 539 TTCTACAAACCCGCTCTTTTACGGAATACAAAGGCACTTAAAGCAACCCCTTTGGAA 598
Db 623 ACATTTGAAGTCACTCTGAGTAAGCAAGTGAACATGCGCTTGAGCAATCTCTCCATA 682
QY 599 AAGTTGCCAAGATAGAGGGCGGCGAGTACATCTCTTCTATCAACAAGATCTCTCA 658
Db 683 GAGCATTTCAAGAGTTGAGACGCTACTCTATCTATCTACGAGAGGAGGAGAACAGAA 742
QY 659 TAACAAGACTTTACTTAACTTGTAAAGTTAGAGTTCAATTTGCTTCACTGCTTGCACAA 718
Db 743 GAATGATGTTGCTTCAATTTGCAAACTGGAATTTCACTTCACTTCACTGCTTGCACAA 802
QY 719 GGAAGAGCTCAGCCATGTTGCAAAATGTTGGAAGCTTTTCGATATCAAGAAGACGACC 778
Db 803 ACAAGAACTTAGTGAAGTATCAAGGTGTTGGAAGATTTTGAATTTGTGACAACTTCC 862
QY 779 TTGTTAAGAGATAGAAATGTTGAATGCTATTTTGGGAGCTAGGTTCAAGGCTATGAGCC 838
Db 863 ATATGCTAGGATAGAGCAGTGAATGCTATTTTGGACGATGGGGGTGTATGCTGAACC 922
QY 839 ACAGTATTCGGGCTAGAGTTTCTTCAAAAGCTGTTGCTGTTTAACTTATATAGA 898
Db 923 TCAATACTCTCAGGCTCTGTCATGCTTGAAGACTATAGCAATGATTTCTATAGTAGA 982
QY 899 TGACACTTATGATGCTGATGTTATGAAAGCTTTTGAAGACTTTTACTGAAGCTTTGA 958
Db 983 TGACACATTCGATGTTATGCAATGTTCAAGAACTTGAGATCTACACCGATGCCATACA 1042
QY 959 AAGTGGTCAATTAATGCTTTAGACACACTTCCAGAAATACATGAACCGATATACAAAT 1018
Db 1043 GAGGTGGGATATTAGCCAAATTTGATCGGCTCCTGATTTACATGAATCAGTTTACAAAG 1102
QY 1019 ATTATGATATACATACAGAAATGGAAGATTTTTCGAAGGAGGAGGAGAGATCT 1078
Db 1103 ACTTTAGATCTCTCAATGATTTGAAATGAGTTGTTCAAGGATGGTAGATCTGAT 1162
QY 1079 ATTTAACTGCGGCAAGAAATTTGTAAGAGAGTTTGTAGAAACCTGATGTTGAAGCAAA 1138

Db 1163 TGTTCACTACGCGAAAGAAATGAAAGAAATCGTGAGAAATATTTTGTGGAGCAAA 1222
QY 1139 ATGGGCAATGAGGAGACATACCAACCACTGAAGAGCATGATCCAGTTGTAATCATTAC 1198
Db 1223 ATGTTTCAATGAAGGATATATGCGGCAGTCTCTGAGTATCTTAGCAATGATTAGCTAC 1282
QY 1199 TGCGGTGCTAACTGCTTCAACCACTGTTATCTTTGGCATGATGATATTTACAAA 1258
Db 1283 CAGCACTTATTTACTTGTCTACGACTACATCTTATTTGGGCATGAATCTGCTAACAGCA 1342
QY 1259 AGAGTCTGCGAATGAGGCTGCTCTGCACTCTCTCTTTTATAGATACCTCAGGTATCTGG 1318
Db 1343 AGATTTTG---AATGTTGGCCAAGAACCCCTAAATTTCTTGAGGCTAATGTGACGTTATG 1399
QY 1319 TCAGCCCTAAATGATCTCATGCCCCCAAGCGGAGCAAGAAAGAAACATAGTTTCATC 1378
Db 1400 CCGAGTCATAGATGACATAGCCACCTATGAGTTGAGAGGTTAGAGTTCAGATTGCCAC 1459
QY 1379 GAGCCTTGAAGTTATATGAAGGAATATAATGTCATGAGGAGTATGCCCAAACTTGTAT 1438
Db 1460 TGGAAATGGAATTTATCATGAGAGATTATGTTGATTCACAGAAAGGCGCATGGAATAAT 1519
QY 1439 TTCAAGAGAGTGAAGATGTTGGAAGATATAAATGTCATGAGGAGTATGCCCAAACTTGTAT 1498
Db 1520 CCAAGAAATGGCTGAGACAGCATGGAAGGATGTAATGAAGAAATCTCTTCGACCACTCC 1579
QY 1499 CATTCAGAGCCGCTTATTTGATGCTGATCTATTTGTCAGTCTTTCTTGAAGTTCAATA 1558
Db 1580 CGTCTCTACAGAGTTCTCATCTGCAATCTCAATCTTCTGCTCGCATTCGATTTACTTTA 1639
QY 1559 TGCAGGAA---AGGATAACTTCACAGTATGGGAGAGCAATPACAAACATCTCATAAAGTC 1615
Db 1640 TAAGCACAATCAAGATGATACACTCATCCGAAAAAGTAGTACTAAAAACCTCATATTATTGC 1699
QY 1616 TCTACTCGTTTATCTTATGATGATATGA 1643
Db 1700 GTTGTGGTGGACTCTATTGAAATTTAA 1727

RESULT 11
US-09-893-820-31
; Sequence 31, Application US/09893820
; Publication No. US20040053386A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20040053386A1, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/893,820
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US/09/398,395A
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 1944
; TYPE: DNA
; ORGANISM: Solanum tuberosum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (57)...(1724)
; OTHER INFORMATION: vetispiradiene synthase
US-09-893-820-31

Query Match 20.0%; Score 330.4; DB 13; Length 1944;
Best Local Similarity 52.0%; Pred. No. 4.5e-78;
Matches 857; Conservative 0; Mismatches 761; Indels 30; Gaps 4;

QY 14 AGAAGAAACCTATTGCGCCCATGCGCAATTCCTCCCAAGCATTTGGGGAGATCAGTT 73
DB 92 AGAGAGGAGATGTTGCGCCCATGCTGCTCTCCAAAGTCTTTGGGGTATCGTTT 151
QY 74 TCTCATCTATCAAAAGCAGTAGAGCAAGGGTGGACAGATGATCAATGATTAATAA 133
DB 152 TCATTCACTTCCCTCGCAATTCAGATTCGCGAAATATGCTCAAGAGATCGAAACTTT 211
QY 134 AGAAGTGGGCACTACTAAAGAACTTTGGGATATTCCTATGAAACATGCCAATTTGTT 193
DB 212 GAAGGAACTCAAGAAATATATATCTGATCTCTCGAAGACATTCGCTGAGAAAT 271
QY 194 GAAGCTGATGATGAATAATCAAGCCCTGGGAATACCGTATCACTTTGAAGCGGAGATGA 253
DB 272 GATCTGATGACATTTGTTGAGCGCTTGGCATTCGCTTATCATTTTGAATAAATAA 331
QY 254 TCATGCTGCAATGATTTATGAACATATGTT-----GATGACTGGAA 298
DB 332 TGATATGTTGGATCAATTTTCAAGCAGATCTTAATCTTTGAGGCTCAGGATCAATGA 391
QY 299 TGGTGACCGCTCTTCTTATGTTTCCGCTCTTATGCGAAAGCAAGGATATTTATGTTACATG 358
DB 392 TTTACAACTTTATCGTTCAATTTTCGACTTATGAGACAACTGTTTACAAATCTCCCC 451
QY 359 TGATGTTTCAATACTATAAGACAAATAATGAGCGGTTCAAGCAATCGTTGCTGATGA 418
DB 452 AAAAATTTTATAGATTCGAAGATCAAAAGGCAAAATTTAAAGAAATCTCTTTGTAACGA 511
QY 419 TCTTGAAGGTTTCTGTTGAGTTGACGAACCACTTCTATGAGGTTACCTGGGAGATTAT 478
DB 512 CATCAAGGCTCTTTGAACTTATACGAAGCTCGCATGTAAGGACTCATGGAGAAATAT 571
QY 479 ATTAGAAGATGCTCTTTGTTTACAGGATCTCTCTTTAGCATTTATGACAAAAGATGCTTT 538
DB 572 TTTGAGAGGACATGCTTTCTTCTACTGCTCATCTTGAATCT-----CGAGCTCC 622
QY 539 TTTACAAACCCCGCTCTTTTACCGAAATACAAAGCGGCACTAAAGCAACCCCTTTGGAA 598
DB 623 ACATTTGAAGTCACTCTGAGTAAGCAAGTGACACATGCGCTTTGAGCAATCTCTCCATAA 682
QY 599 AAGGTTGCCAAGATAGAGCGCGCAGTACATCTCTTCTATCAACAAAGATTTCTCA 658
DB 683 GAGCATTTCAAGATTTGAGACAGCTACTCTCTCTATCTAGCAAGAGGAGAAACAGAA 742
QY 659 TAACAGACTTTTACTTAAACTTCTTAAGTTAGAGTTCAATTTGCTTTCAGTCAATTGCACAA 718
DB 743 GAATGATGTTGCTTCAATTTGCAAACTGGACTTCACTTACTTACAGATTTGCACAA 802
QY 719 GGAAGAGCTCAGCCATGTTGCAAAATGGTGGAAAGCTTTTCGATATCAAGAAAGCAGCACC 778
DB 803 ACAAGAACTTAGTGAAGTATCAAGGTGGTGGAAAGATTTGGATTTTGTGACAACTTCC 862
QY 779 TTGTTTAAGAGATAGATTTCTTGAAGTCTACTTTTGGGACTAGGTTTTCAGGCTATGAGCC 838
DB 863 ATATGCTAGGATAGAGCAGTGAATGCTACTTTTGGACGATGGGGGTGATGCTGAACC 922
QY 839 ACAGATTTCCCGGCTAGAGTTTCTTTCACAAAGCTTTGCTGTTTATTAATCTTTATAGA 898
DB 923 TCAATACTCTCAGCTCGTGTGATGCTGCTAGACTATAGCAATGATTTCTATAGTAGA 982
QY 899 TGACACTTATGATGCTATGTTGATCTTATGAGAACTTAAAGATCTTTACTGAAAGTGTGA 958
DB 983 TGACATTCGATGCTTATGGCATTTGCAAGAACTTGGATCTACACCGATGCCATACA 1042
QY 959 AAGGTGGTCAATFACATGCTTACACACACTTCCAGATATACATGAACCGATATACAAATT 1018
DB 1043 GAGGTGGATATTAGCCAAATGATCGGCTCCCTGATTTACATGAAATCAGTTACAAAGC 1102
QY 1019 ATTCTGGATACATACACAGAAATGGAAGAAATTTCTTCCAAAGAGGGAAGAAACAGATCT 1078

DB 1103 ACITTTAGATCTCTACAAATGATTATGAATGAGTTGTCGAAGGATGGTAGATCTGATGT 1162
QY 1079 ATTAACTGCGGCAAGAAATTTGTTGAAAGAGTTTGTAGAAACCTGATGGTTGAAGCAA 1138
DB 1163 TGTTCACCTGCGGAAGAAAGAAATCGTGAGAAACTATTTTGTGGAAGCAA 1222
QY 1139 ATGGGCAATGAGGACACATACCAACCACTGAGAGCATGATCCAGTTGTAAATCAATTAC 1198
DB 1223 ATGTTTCAATTTGAAGGATATATGCGCCAGTCTCTGAGTATCTTAGCAATGCAATAGCTAC 1282
QY 1199 TGGCGGTGCTAACTGCTTACAAACAACTTTGTTATCTTGGCATGAGTGTATATTCACAA 1258
DB 1283 CAGCACTTATTTACTTGTCTTACGACTACATCTTATTTGGGCATGAAAGTCTGCTAAACAAGCA 1342
QY 1259 AGACTCTGCGAATGGGCTGCTCTGACACCTCTCTTTTATAGATACTCAGGTATACTTGG 1318
DB 1343 AGATTTTG---AATGGTTGGCCAAAGACCCCTTAAATTTCTTGAGGCTAATGTGACGTTATG 1399
QY 1319 TCGACGCTCTAAATGATCTCATGACCCCAAGGCGGAGCAAGAAAGAAACCAATAGTTTCATC 1378
DB 1400 CCGAGTCAATAGATGACATAGCCACTATGAGGTTGAGAAGGGTAGAGGTCAGATTGCCAC 1459
QY 1379 GAGCTTTGAAAGTTATATGAAGGAATATAATGTGCAATGAGGAGTATGCCAAACCTTGAT 1438
DB 1460 TGGAAATGAAATTTTACATGAGAGATTTATGTTGTATCCACAGAAAGGCCATGGAATAAT 1519
QY 1439 TTCAAGGAGTGAAGATGTTGCGAAGATATAAACCAGAGATACCTCACAACTAAATA 1498
DB 1520 CCAAGAAATGCTGAGACAGCATGGAAGGATGTAATGAAGGAATCCTTCGACCAACTCC 1579
QY 1499 CATTCGAAGGCGGCTTATTTGATGGCTGATCTATTTGTCGCAAGTTTCTTGAAGTTCAATA 1558
DB 1580 CGTCTCTACAGAGATTTCTCACTCGCATCTCAATCTTCTGCTCGCATTCGATGTTACTTA 1639
QY 1559 TGCAGGAA---AGGATACTTCAACAGTATGGGAGAGCAATACAAACATCTCATAAAGTC 1615
DB 1640 TAAGCAAAATCAAGATGATACATCTCATCCGAAAGAGTACTTAAACCTCATATATTATTCG 1699
QY 1616 TCTACTCGTTTATCTCTATGATATGA 1643
DB 1700 GTTCTGCTGACTCTATTTGAATTTAA 1727

RESULT 12
US-09-887-586A-1
; Sequence 1, Application US/09887586A
; Patent No. US20020094556A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20020094556A1, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/887,586A
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1671
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (25)...(1668)
US-09-887-586A-1

Query Match 19.4%; Score 320.6; DB 9; Length 1671;
Best Local Similarity 52.0%; Pred. No. 1.8e-75;
Matches 850; Conservative 0; Mismatches 764; Indels 21; Gaps 5;

15 GAAGAAAACCTATTTCGCCCATTCGCCCATTCCTCCCAAGCATTTGGGAGATCAGTTT 74
Db |||||
52 GAAGAAGAGATTGTTGCCCGCGTCCGCCACTTCTCCCTAGTTTGTGGGGTGATCAGTTT 111
Qy |||||
75 CTATCTATCAAAAGCAAGTAGAGCAAGGGTGGAAACAGATAGTGAATGATTTAAAAAAA 134
Db |||||
112 CTTTCATCTCCATTA---AAATCAGGTTGCAGAAAGATAGTCTCAAGAGATTGAACA 168
Qy |||||
135 GAAGTGGCGCAACTACTAAAGAAAGCTTTGGATATTCCTATGAACATGCCCAATTTGTTG 194
Db |||||
169 TTGAAGGAACAACGAGGAAATATGCTGTTAGCAACTGGAATGAAATTTGCTGATACACTG 228
Qy |||||
195 AAGCTGATGATGAATTAACGCGCTTGGAAATACCGTATACATTTGAACGGGAGATTGAT 254
Db |||||
229 AATTGTATGACACTATGTAACGCGCTTGGCATATCTTACCACCTTTTGAGAAAGAAATTGAT 288
Qy |||||
255 CATGCATTGCAATGATTTATGAACATATGTTGATTAACCTGGAATGGTGACCGCTCTTCC 314
Db |||||
289 GATATTTTGGATCAGATTTTACAACAACTCAAACTGCAAGATTTGTGCACTTCTGCA 348
Qy |||||
315 TTATGTTTCCGCTCTATGCGAAAGCAAGGATATTTATGTTACATGTGATGTTTTCATTAAC 374
Db |||||
349 CTTCAATTTCCGATGCTCAGGCAACATGTTTCAACATCTCTCCGAAATTTTCAGCAA 408
Qy |||||
375 TATAAGACAAAATAGGAGCGCTTCAAGCAATCGTTAGCTAATGATGTTGAAGTTTGCTT 434
Db |||||
409 TTCCAGACGAAATGGAATTTCAAGGATCTCTTGTCTAGTATGTTCTTAGGATTTATG 468
Qy |||||
435 GAGTTGTACGAAGCAACTTCTATGAGGCTACTGGGAGATATATATTAAGAGATCCTTT 494
Db |||||
469 AACTTGTATGAAGCTTACATGTAAGGACTCATGCTGAGATCTTTTGAAGACGCATTT 528
Qy |||||
495 GGTTTACACGATCTCGTCTTAGCATTTATGACAAAAGATGCTTTTCTACAAACCCCGT 554
Db |||||
529 GCTTTCTC-----CACTATCCATCTGAACTCTGAGCTCCACATTTGAAATCTCA 579
Qy |||||
555 CTTTTCACGAAATACAAACGGGCACTAAGCAACCCCTTTTGGAAAGGTTGCAAGATA 614
Db |||||
580 CTTAGGGAGCAAGTGACACATGCGCCTTGAGCAATGTTTGCAACAGGGTGTCTTAGAGTC 639
Qy |||||
615 GAGCGCGCGAGTACATTCCTT---TCTATCAACACAGATTTCTATCAACAGACTTTA 671
Db |||||
640 GAGACCGGATTTCTCATCTCATCAATCTATGACAAAGGAACATCGAAGATATGTTGTTA 699
Qy |||||
672 CTTTAAACCTTGCTAAGTTAGAGTTCAATTTGCTTCAGTCAITGCAACAAAGAGCTCAGC 731
Db |||||
700 CTTGCAATTTGCCAATTTGATTTCACTTGTCTCAGATGTTGCAACAAAGAACTTGTCT 759
Qy |||||
732 CATGTGCAAAATGGTGGAAAGCTTTTCGATATCAAGAAAGACGCACTTTGTTTAAAGAT 791
Db |||||
760 CAAGTATCAAGGTGGTGGAAAGATTTGGAATTTGTAACAACACTTCCATATGCTAGAGAT 819
Qy |||||
792 AGAATTTGGAATGCTACTTTTGGGACTAGTTCAGGCTATGAGCCACAGTATTTCCCGG 851
Db |||||
820 CGAGTAGTTGAATGCTACTTTTGGGCATTAGAGTTTATTTTGAGCCCTCACTCTCAA 879
Qy |||||
852 GCTAGAGTTTCTTCAAAAAGCTTGTCTGTTTAACTCTTATAGATGACACTTATGAT 911
Db |||||
880 GCTCGGCTATGCTCGTTAAGACCATACTAATGATTTGATTTGATGACACCTTTGAT 939
Qy |||||
912 GCGTATGCTATTATGAAGAACTTAAGATCTTTACTGAAGCTGTTTGAAGGTTGCAATT 971
Db |||||
940 GCTTACGGTACAGTTAAGAACTTGAGGCATACACAGATGCGCATACAAAGATGGGATATC 999
Qy |||||
972 ACATGCTTAGACACATTCAGAAATACATGAACCGGATATCAATATTTATGATGATACA 1031
Db |||||
1000 AACGAAATGATCGGCTTCTCGATTACATGAAATCAGTTTCAAAAGCTATTCTAGATCTC 1059

1032 TACACAGAAATGGAAGAAATTTCTTGGCAAGGAGGAAAGACAGATCTATTTTAACTCGCGC 1091
Db |||||
1060 TACAGGATTATGAAAGGAATTTGTCTAGTGGCGGAAGATCTCATATTTGTCTGCCATGCA 1119
Qy |||||
1092 AAGAAATTTGTGAAGAGTTTGTAGAAACCTGATGTTGAGCAAAATGCGCAATGAG 1151
Db |||||
1120 ATAGAAGAATGAAGAGTAGTAAGAAATTTAATGTCGAGTCAACATGTTTATTGAA 1179
Qy |||||
1152 GGACACATACCAACCACTGAAGAGAGATGATCCAGTTGTAATCAATTAATCTGGGGTCTTAAC 1211
Db |||||
1180 GATATACGCCACTGTTTCTGATACCTAAGCAATGCACTAGCAACTACACATATTAC 1239
Qy |||||
1212 CTGCTTACAAACAACTGTTTATCTTGGCATGAGTGATATTTCAAAAAGAGTCTGTCGAA 1271
Db |||||
1240 TACCTCGACACACATCGTATTTTGGGCATGAAGTCTGC---CACGGAGCAAGATTTTGAG 1296
Qy |||||
1272 TGGGCTGTCTCTGCACCTCTCTTTTAGATATCTCAGGTATACTTGGTCGAGCCCTAAAT 1331
Db |||||
1297 TGGTTGTCAAAGAAATCCAAAAATTTCTTGAAGCTAGTGTAAATTTATGTCGAGTTATCGAT 1356
Qy |||||
1332 GATCTCATGACCCACCAAGCGCGGAGCAAGAAAGAAAACATAGTTTCATCGAGCCTTGAAAAGT 1391
Db |||||
1357 GACACAGCCACGTACGAGGTTGAGAAAGCAGGGGACAAATTTGCACTGGAATTTAGTGC 1416
Qy |||||
1392 TATATGAAGAAATATAATGTCAATGAGGAGTATGCCCAACCTTGAATTTACAAGGAAGTA 1451
Db |||||
1417 TGCATGAGAGATTATGATATATCAACAAAGAGGCAATGGCTAAATTTCAAAATATGGCT 1476
Qy |||||
1452 GAGATGTGTGAAAGATATAAACCAGAGAGTACTCAACAATAAAACATTTCCAGGCCG 1511
Db |||||
1477 GAGACAGCATGGAAGATATTAATGAAGGACTTCTTTAGGCCCACTCCCGTCTCTACAGAA 1536
Qy |||||
1512 TTATTCATGCTGTGATCTATTTTGTGCGAGTTCTTGAAGTTCAATATGCAAGAAA---G 1568
Db |||||
1537 TTTTAACTTCTTATTTCTCAATCTTGTCTGATTTGTTGAGTTATATATACATCTA 1596
Qy |||||
1569 GATACTTCAACGATGAGGAGAGCAATAACAACATCTCAATAAGTCTCTACTCGTTTAT 1628
Db |||||
1597 GATGGATACACTCATCCGGAGAAAGTCTTAAACCTCACATTTATTAACCTACTTTGGGAC 1656
Qy |||||
1629 CCTATGAGTATAGA 1643
Db |||||
1657 TCCATCAAAATTTGA 1671

RESULT 13
US-09-903-012-1
; Sequence 1, Application US/09903012
; Patent No. US20020094557A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20020094557A1, Joseph P.
; APPLICANT: Starke, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASE
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/903,012
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1671
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
; FEATURE:

NAME/KEY: CDS
LOCATION: (25)...(1668)
US-09-903-012-1

Query Match 19.4%; Score 320.6; DB 9; Length 1671;
Best Local Similarity 52.0%; Pred. No. 1.8e-75;
Matches 850; Conservative 0; Mismatches 764; Indels 21; Gaps 5;

15 GAAGAAAACCTATTTCGCCCATTCGCCAATTCCTCCCAAGCATTTGGGGAGATCAGTTT 74
DB GAAGAGAGATTTGGCCCGCTGCCGACTTCCTCCCTAGTTTGGGGGATCAGTTT 111
75 CTCATCTATCAAAAGCAAGTAGAGCAAGGGTGGAAACAGATAGTGAATGATTTAAAAAAA 134
DB CTTTCATCTCCATTA--AAATCAGGTTGCAGAAAAGTATGCTCAAGAGATTGAAGCA 168
135 GAAGTGGCGCACTACTAAAGAGCTTTGGATATTCCTATGAACATGCCAATTTGTTG 194
DB 169 TTGAAGGACAAACGAGGAATATGCTGTAGCAACTGGAATGAAATTTGGCTGATACACTG 228
195 AAGCTGATTGATGAATTAACAGCTTGGAAATACGATATCACTTTGAACGGGAGATTGAT 254
DB 229 AATTGTAGACACTATTGAACGCTTGGCAATCTCTACCATTTTGAAGAAATGAT 288
255 CATGCAATTGCAATGATTTATGAACATATGGTGATTAACCTGGAATGGTGACCGCTCTCC 314
DB 289 GATATTTTGGATCAGATTTTACCAACCAAACTCAAACTGCAACGATTTGTGCACTTCTGCA 348
315 TTATGGTTCGCTTATGCGAAGCAAGGATATTTATGTTACATGATGATGTTTCAATAAC 374
DB 349 CTTCAATTTTCGATTCCTCAGGCAACATGGTTCCTCAACATCTCTCTGAAATTTTCAGCAA 408
375 TATAAGACAAAATGGAGCGTTCAAGCAATCGTTAGCTTAATGATGTTGAAGTTTGGCTT 434
DB 409 TTCCAAGACGAAATGGCAATTTCAAGAACTCTTGTGATGATGCTTAGGATTTTG 468
435 GAGTTGTACGAAGCAACTTCTATGAGGTPACCTGGGGAGATATATTAAGAAGTCTCTT 494
DB 469 AACTTGTATGAAGCTTCACATGTAAGGACTCATGCTGACGATATCTTAGAAGCGCACTT 528
495 GGTTTTACAGATCTGCTTAGCATTTATGCAAAAGATGCTTTTCTCAAAACCCGCT 554
DB 529 GCTTTCTC-----CACTATCCATCTTGAATCTGACGCTCCACATTTGAATCTCCA 579
555 CTTTTCACGAATAACACGGGCACCTAAAGCAACCCCTTTGGAAAAGGTTGCCAAGAAATA 614
DB 580 CTTAGGGAGCAAGTACACATGCCCTTGAGCAATGTTTGCAACAGGCTGTTCTAGATC 639
615 GAGCGCGCAGTACATTCCTT---TCTATCAACAACAAGATTCCTCAACAAGACTTTA 671
DB 640 GAGACCCGATTCCTCATCTCATCAATCTATGCAAGGAACAATCGAAGAAATATGTTTA 699
672 CTTAACTTGTCTAAGTTAGATTCATTTGCTTCAGTCAITGCAACAGGAGAGCTCAGC 731
DB 700 CTTGCAATTTGCCAAATTTGATTTCAACTGCTCCAGATGTTGCAACAAACAAGAACTTGT 759
732 CATGTGTGCAATTTGGTGAAGCTTTTCGATATCAAGAAGAACGACCTGTTTAAAGAGAT 791
DB 760 CAAGTATCAAGTGTGGAAGATTTGATTTGTAACAACACTTCATATGCTAGAT 819
792 AGAATTTGTAATGCTACTTTTGGGAGTGTGCTCAGGCTATGAGCCACAGTATTTCCCGG 851
DB 820 CGAGTAGTTGATGCTACTTTTGGGCATTAGGAGTTTATTTTGGAGCTCAATCTCTCAA 879
852 GCTAGATTTTCTTCAAAAGCTTGTGTTTATTAACCTTTATAGATGACATTTATGAT 911
DB 880 GCTCGCTCATGCTCGTTAAGACCATATCAATGATTTTCGATTCGATGACACTTTGAT 939
912 GCGTATGTTACTTATGAAGAACTTAAGATCTTTTACTGAAAGCTGTTGAAGGCTGCTCAATT 971
DB 940 GCTTACGGTACAGTTAAAGACTTTAGGCATACACAGATGCCATACAAAGATGGGATTC 999
972 ACATGCTTAGACACACTTCCAGAAATACATGAACCGGATATACAAATTTATTCATGATACA 1031

1000 AACGAAATTGATCGGCTTCCTGATTACATGAAATCAGTTTACAAAGCTATTCTAGATCTC 1059
QY 1032 TACACAGAAATGGAAGATTTCTTGCAGAGGGAAGAGACAGATCTATTTAACTGCGGC 1091
DB 1060 TACAAGGATTTATGAAGGAATTTCTTAGTCCCGAAGATCTCATATTTGTCTGCATGCA 1119
QY 1092 AAAGAATTTGTGAAGAGTTTGTAGAAAACCTGATGTTGAAGCAAAATGGGCAAAATGAG 1151
DB 1120 ATAGAAGAATGAAGAGTAGTAAGAAATTAATTCGAGTCAACATGGTTTATTGAA 1179
QY 1152 GGACACATACCAACCACTGAAGAGCATGATCCAGTTGTAATCATTTCTGGCGGTCTAAC 1211
DB 1180 GGATATACGCCACCTGTTCTGAATACCTAAGCAATGCACTAGCAACTTACCACATATTAC 1239
QY 1212 CTGCTTACAACTGTTTATCTTGGCATGATGATATATTCACAAAGAGCTCTGCGAA 1271
DB 1240 TACCTCGGCAACATCGTATTTGGGCATGAAGTCTGC---CACGGACAAAGTTTGGAG 1296
QY 1272 TGGGCTGTCTCTGCACCTCTCTTTTGTAGATACCTCAGTATACCTTGGTTCGACGCTTAAAT 1331
DB 1297 TGGTTGTCAAGAAATCCAAAATTTCTTGAAGCTAGTGTAAATATATGTCGATTCGAT 1356
QY 1332 GATCTCATGACCCCAAGGCGGAGCAAGAAAGAAACATGTTTCAATTCGAGCCTTGAAAGT 1391
DB 1357 GACACAGCCACGTACGAGGTTGAGAAAAGCAGGGGACAAATTTGCAACTGGAATTTGAGTGC 1416
QY 1392 TATATGAGGAATATATGTCMAATGAGGAGTATGCCCAACCTTGATTTACAAGGAAGTA 1451
DB 1417 TGCATGAGATTTATGATATATCAACAAAGAGGCAATGGCTAAATTTCAAAATATGGCT 1476
QY 1452 GAAGATGTTGGAAAAGATATAAACCAGAGAGTACCTCAACATAAAACATTTCCAAAGCGCG 1511
DB 1477 GAGACAGCATGGAAGATATTAATGAAGGACTTCTTAGGCCCACTCCCGTCTCTACAGAA 1536
QY 1512 TTATGATGCTGTGATCTATTTGTGCGAGTTTCTTGAAGTTCAATATGCGAGAAA---G 1568
DB 1537 TTTTAACTCCTATCTCAATCTTGTCTGCTGATTTGAGGTTTACATATATACAAATCTA 1596
QY 1569 GATAACTTCACGCTATGGGAGAGCAATACAAACATCTCAATAAGTCTCTACTCGTTTAT 1628
DB 1597 GATGATACACTCATCCGAGAAAGTCTTAAACCTCACATTTAACTTAACTTCTGTGGAC 1656
QY 1629 CCTATGATATATGA 1643
DB 1657 TCCATCAAAATTTGA 1671

RESULT 14
US-09-900-797-1
Sequence 1, Application US/09900797
Publication No. US20030087406A1
GENERAL INFORMATION:
APPLICANT: Chappell, Joseph
APPLICANT: No. US20030087406A1, Joseph P.
APPLICANT: Starks, Courtney M.
APPLICANT: Manna, Kathleen R.
TITLE OF INVENTION: SYNTHASES
FILE REFERENCE: 07678-025001
CURRENT APPLICATION NUMBER: US/09/900,797
CURRENT FILING DATE: 2001-07-06
PRIOR APPLICATION NUMBER: US/09/398,395
PRIOR FILING DATE: 1999-09-17
PRIOR APPLICATION NUMBER: 60/130,628
PRIOR FILING DATE: 1999-04-22
PRIOR APPLICATION NUMBER: 60/150,262
PRIOR FILING DATE: 1999-08-23
NUMBER OF SEQ ID NOS: 58
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 1671
TYPE: DNA
ORGANISM: Nicotiana tabacum

```
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (25)....(1668)
US-09-900-797-1

Query Match      19.4%; Score 320.6; DB 10; Length 1671;
Best Local Similarity 52.0%; Pred. No. 1.8e-75;
Matches 850; Conservative 0; Mismatches 764; Indels 21; Gaps 5;

Qy 15 GAAGAAAACCTTATTCGCCCAATTCGCAATTCCTCCCAAGCAATTTGGGAGATCAGTTT 74
Db 52 GAAGAGAGATTGTCGCCCGCTCGCGGACTTCTCCCTAGTTTGGGGTGAICAGTTT 111
Qy 75 CTCTCTATCAAAAGCAAGTAGAGCAAGGGGTGGAAACAGATAGTGAATGATTTAAAAAAA 134
Db 112 CTCTCTATTCCTCAATTA ---AAATCAGGTTGACAGAAAAGTATGCTCAAGAGATTGAACA 168
Qy 135 GAAGTGGCGCACTACTAAAAGAGCTTTGGATATTCCTATGAAAACATGCCAATTTGTTG 194
Db 169 TTGAAGGAACAAAACGAGGAATATGCTTTAGCAACTGGAATGAAATTTGGCTGATACACTG 228
Qy 195 AAGCTGATTGATGAATTCACCGCTTGAATACCGTATCCTTGAACGGGAGATTGAT 254
Db 229 AATTTGATAGACACTATTGAAACGCTTTGGCAATCTCTACACTTTGAGAAGAAATTTGAT 288
Qy 255 CATGCAATTGCAATGATTTATGAACATATGTTGATACTGGAATGGTGACCCGCTCTTCC 314
Db 289 GATATTTTGGATCAGATTTTACACCAAACTCAAACTGCAACGATTTGTGCACCTTCTGCA 348
Qy 315 TTATCGTCCGCTTATGCGAAGAGAGATATTTATGTTACATGATGATTTTCAATAAAC 374
Db 349 CTTCATTTTCGATTGCTCAGGCAACATGTTTCAACATCTCTCCTGAAATTTTTCAGCAA 408
Qy 375 TATAAGACAAAATTTGAGCGTTTCAAGCAATGTTAGCTAAATGATGTTGAAGGTTTGCCT 434
Db 409 TTCCAAGACGAAAATTTGCGAAATTTCAAGGAATCTTCTGCTAGTATGCTTTAGGATTTTG 468
Qy 435 GAGTTGTAGCAAGCACTTCTATGAGGGTACCTGGGGAGATATATTTAGAAAGTCTCTT 494
Db 469 AACTTTGATGAGCTTCAATGTAAGGACTCATGCTGACGATATCTTTAGAAAGCGCACT 528
Qy 495 GGTTTTACAGATCTGCTTTAGCATATGACAAAAGATGCTTTTCTACAAACCCGCT 554
Db 529 GCTTTCTC-----CACTATCCATCTTGAATCTGCAGCTCCACATTTGAAATCTCCA 579
Qy 555 CTTTTTACCGAATACACGGGCACTTAAGCAACCCCTTTGGAAGAGGTTGCCAAGATA 614
Db 580 CTTAGGAGCAAGTGACATGCCCTTTGAGCAATGTTTGCACAAAGGTTTCTCTAGATC 639
Qy 615 GAGCGGGCGAGTACATTCCTT---TCTATCAACAAGATTTCTCATACAAGACTTTA 671
Db 640 GAGACCCGATTCTTCATCTCATCAATCTATGACAGGAACAATCGAAGATAATGTGTTA 699
Qy 672 CTTAAACTTGTAAGTTAGATTCATTTGCTTCACTGCTATGACAAAGGAGAGCTCAGC 731
Db 700 CTTCCGATTGCAAAATTTGATTTCACTTGCCTCAGATGTTGCACAAACAAGAACTTGCT 759
Qy 732 CATGTTGCAATGTGGAAAGCTTTCGATATCAAGAAGACGCACTTGTGTTTAAAGAT 791
Db 760 CAAGTATCAAGTGTGGAAAGATTGGATTTTGTAAACAACACTTCCATATGCTAGAT 819
Qy 792 AGAATGTTGATGCTACTTTTGGGACTAGTGTTCAGGCTATGAGCCACAGTATTTCCCGG 851
Db 820 CGAGTAGTTGATGCTACTTTTGGGCAATAGGAGTTTATTTGAGCCTCAATCTCTCA 879
Qy 852 GCTAGAGTTTCTTCAAAAAGCTGTTGCTGTTATTAATCTCTATAGATGACATTTATGAT 911
Db 880 GCTCGCTCATGCTCGTTAAGACCATATCAATGATTTTCGATTTGCTGATGACACCTTTGAT 939
Qy 912 GCGTATGTTACTTATGAAGCTTATAGATCTTATGAGGCTGTTGAAAGGTTGCTCAATT 971
Db 940 GCTTACGGTATAGTTAAAGAACTTTAGGGCATACACAGATGCCATACAAAGATGGGATC 999
```

RESULT 15

```
US-09-893-820-1
; Sequence 1, Application US/09893820
; Publication No. US20040053386A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20040053386A1, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/893,820
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US/09/398,395A
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
```

```

; LENGTH: 1671
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (25)...(1668)
US-09-893-820-1

Query Match      19.4%; Score 320.6; DB 13; Length 1671;
Best Local Similarity 52.0%; Pred. No. 1.8e-75;
Matches 850; Conservative 0; Mismatches 764; Indels 21; Gaps 5;

QY 15 GAAGAAAACTATTGCGCCCATTTGCCAACTTTCCTCCAGCACTTTGGGAGATCAGTTT 74
Db 52 GAAGAAGAGATGTTTCGCCCGCGTCGCCGACTTCCCTCTAGTTTGTGGGTGATCAGTTC 111
QY 75 CTCATCTATCAAAAGCAAGTAGAGCAAGGGGTGAAACAGATAGTGAATGATTAAAAAAA 134
Db 112 CTTTCATTCTCCATTAA---AAATCAGGTTGCAGAAAAGTATGCTCAGAGATTGAAGCA 168
QY 135 GAAGTCGGCAACTACTAAAGAAGCTTTGGATATTCTATGAACAATGCCAATTTGTTG 194
Db 169 TTGAAGGAACAACACGAGGAATATGCTGTGTAGCAACTGGAATGAAATTTGGCTGATACACTG 228
QY 195 AAGCTGATTGATGAATTCACGCCCTTGGATACCGTATCACITTTGAACGGGAGATTGAT 254
Db 229 AATTGTGATAGACACTATTGAACGCCCTTGGCATATCCTACCACCTTTGAGAAGAAATTTGAT 288
QY 255 CATGCATTGCAATGATTATTATGAACAATATGGTGTAACTTGAATGGTGACCGCTCTTCC 314
Db 289 GATATTTTGGATCAGATTACAAACCAAACTCAAACTGCAACGATTTGTGCACCTTCTGCA 348
QY 315 TTATGGTTCGTTTATTCGGAAGCAAGGATATTATGTTACATGTGATGTTTCAATPAAC 374
Db 349 CTTCAATTTTCGATTGCTCAGGCAACATGGTTTCAACACTCTCCTCGAAATTTTCAGCAA 408
QY 375 TATAAAGACAAAATATGAGCGGTTCAAGCAATCGTTAGCTAATGATGTTAGAGTTTCCTT 434
Db 409 TTCAAGACGAAAATGCGAAATTCAGGAATCTCTTCTAGTGTGATGTTCTTAGGATTAATG 468
QY 435 GAGTTGTACGAAGCAACTTCTATCAGGGTACTCGGGGAGATTAATTATAGAAGATGCTCTT 494
Db 469 AACTTGTATGAAGCTTCACATGTGAAGACTCACTGCTGACGATATCTTTAGAAGCGCACTT 528
QY 495 GGTTTTACCGATCTCGTCTTTAGCATTTATGACAAAAGATGCTTTTCTACAAACCCCGCT 554
Db 529 GCTTTCTC-----CACTATCCATCTTGAATCTCGAGCTCCACATTTGAAATCTCCA 579
QY 555 CTTTTTACCGAAATACAAAGGGGCACTAAAGCAACCCCTTTGGAAAAAGGTTGCCAAGAATA 614
Db 580 CTTAGGGAGCAAGTGACACATGCGCTTTGAGCAATTTTGTGACAAAGGGTGTCTCTAGAGTC 639
QY 615 GAGCGGGCGAGTACATTCCTT---TCTATCAACAACAAGATTTCTCATAAACAGACTTTA 671
Db 640 GAGACCGGATTCCTTCACTCATCTCAATCTATGCAAGGAACAATCGAAGAAATATATGTGTTA 699
QY 672 CTTAAACTTGTCTAAGTTAGAGTTCAATTTTGTCTTCAGTCAATTGCAACAAGGAGAGCTCAGC 731
Db 700 CTTTCGATTTGCCAATTTGGAATTTCAAATTGCTCCAGATGTTGACAAACAAGAACTTGTCT 759
QY 732 CATGTGTGCAAAATGGTGGAAAGCTTTTCGATATCAAGAAAGAACGCACTTGTGTTAAGAGAT 791
Db 760 CAAGTATCAAGGCTGGTGGAAAGATTTTGGATTTTGTAAACAACACATCTCCATATGCTAGAGAT 819
QY 792 AGAATTGTTGAATGCTACTTTTGGGACTAGGTTTCAGGCTATGAGCCACACAGTATTTCCCGG 851
Db 820 CGAGTAGTTGAATGCTACTTTTGGGCATTTAGAGGTTTATTTTGTAGGCTCAATATCTCTCAA 879
QY 852 GCTAGAGTTTTCCTTCAAAAAAGCTGTTGCTGCTTATACTCTTATAGATGACACTTATGAT 911
Db 880 GCTCGGCTCATGCTCGTTTAAAGCCATATCAATGATTTTCGATTTGATGATGACACCTTTGAT 939
QY 912 CGGTATGGTACTTATGAAGAACTTTAAGATCTTTTACTGAAGCTGTGTGAAGGGGTGCAATT 971

```

QY 615 GAGCGGCGCAGTACATTCCTT---TCTATCAACAAGATTCTCATACAAGACTTTA 671

Fax: 1-(530)-752-9659
Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
belongs to contig QG_CA_Contig2363, see http://cgdb.ucdavis.edu/
for details.
Plate: QGC27 row: H column: 08.
Location/Qualifiers
1. 749
/organism="Lactuca sativa"
/mol_type="mRNA"
/cultivar="Salinas"
/db_xref="taxon:4236"
/clone="QGC27H08"
/lab_host="E.coli"
/clone_lib="QG ABCDI lettuce salinas"
/notes="Vector: pBRCDNASFIAB; The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgdb.ucdavis.edu/
TAG_TISSUE="chemical induction"
TAG_LIB="QG ABCDI lettuce salinas"
TAG_SEQ="TGTCAGCCGGG"

ORIGIN

Query Match 23.3%; Score 384.2; DB 13; Length 749;
Best Local Similarity 70.3%; Pred. No. 4.2e-86;
Matches 515; Conservative 0; Mismatches 218; Indels 0; Gaps 0;

QY 109 AACAGATAGTGAATGATTAAAGGAGGCGGCACTACTAAAGAGCTTTGGATA 168
DB 17 ATCGAATAGTCCATGATCTGAAAGAAAGTGAGAAAGATATCTGCTGCTTTATG 76
QY 169 TTCTATGAAACATGCCAATTTGTTGAAGCTGATGATGAATTCACGCTTGAATAC 228
DB 77 TTCCAAATGGAACATACAAAACCTCTTGAACATTAATCGACACATCCAGCTCTTGAATAG 136
QY 229 CGTATCACTTTGACGGGAGATGATCATGATGCAATGATTTATTAAGACATATG 288
DB 137 CCTACTATTTTCGACCAAGAGATTAAGCAAGCCTTCGACATATCTATGTGAATATG 196
QY 289 ATAACTGAAATGGTGACCGCTCTTCCTTATGTTCCGCTTATGCGAAGCAAGGATAT 348
DB 197 ATAACTGAGTGGTGGTGTAGCTCTTCTGTTTGGCTCATGCGACACAGGTTCT 256
QY 349 ATGTTACATGATGTTTTCATTAATACTATAAGACAAAATGAGCGGTTCAAGCAATCG 408
DB 257 ACGTTTCATGATGATTTTCAATACTACAAAAAAGAAATGGAGCTTTTAAGGAATCG 316
QY 409 TAGCTAATGATGTTGAAGGTTGCTTGTGATGTTGACGAGCACTCTTATGAGGTTACCT 468
DB 317 TGACCAATGCATCCATGAGATGCTTGATGTTATATGAGGCAACATTTATGAGAGTGA 376
QY 469 GGGAGATATATTAGAAGATGCTCTTGTTTTCACAGATCTCGTCTTACATTTAGCAA 528
DB 377 GAGAGTTGTACTAGACGAGCTCTTCTTTTCAAAAACATCTTTGAAAACCTAGCAA 436
QY 529 AAGATGCTTTTTCACAAACCCCGCTCTTTTTCACGAAATACAAACGGGCACTAAAGCAAC 588
DB 437 AGGATCCCTGTTCCGTTAACTCTACTCTTCAATACATACAGGAGACACTAAAAACATC 496
QY 589 CCCTTTGAAGAGTTGCCAAGATAGAGGCGCGCAGTACATCTCTTCTATCAACAC 648
DB 497 CCATATTGAGAAGGTTGCCAAGCTTGAAGCTTGAAGCAATACATCTCTTCTACCAAAAC 556
QY 649 AAGATTCCTATAACAGACTTTTACTTAACTTGCCTAAGTTAGAGTTCAATTTGCTTCAGT 708
DB 557 AAGCTTCTTGTAAGTCTTACTTCACTTTTCAAGTTAGGTTCACTTCACTTCAAT 616
QY 709 CATTGCACAAAGAGAGTTCAGCCATGTTGTGCAATGTTGGAAAGCTTTTCGATATCAAGA 768

DB 617 CCCTACACAAAGAGAGCTTAGTGAACCTTCCAGTGGTGAAGGTTTGTGTACCGA 676
QY 769 AGAAGCGACCTTGTGTTAAGAGATAGAAATGTTGAATGCTACTTTTGGGACTAGGTTT 828
DB 677 ACAAGATACCTTATACACGGGACCGATTGTTGAATGTTACTTTTGGGCATTAGTGT 736
QY 829 GCTATGAGCCACA 841
DB 737 ACTTGTAGCCACA 749

RESULT 2
BQ861631
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Cichorioideae;
Cichorieae; Lactuca.
1 (bases 1 to 692)
Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Riesberg, L.,
Lin, H., van Damme, M., Lavelle, D., Chevalier, P., Ziegler, J.,
Ellison, P., Kolkman, J., Slabaugh, M.S., Livingston, K., Zhou, Y.,
Lai, Z., Church, S., Jackson, L. and Bradford, K.
Lettuce and Sunflower ESTs from the Compositae Genome Project
http://comgenomics.ucdavis.edu/
Unpublished (2002)
Contact: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Asmundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@atgc.org [michelmore@vegmail.ucdavis.edu]
belongs to contig QG_CA_Contig2363, see http://cgdb.ucdavis.edu/
for details.
Plate: QGC19 row: C column: 16.
Location/Qualifiers
1. 692
/organism="Lactuca sativa"
/mol_type="mRNA"
/cultivar="Salinas"
/db_xref="taxon:4236"
/clone="QG19C16"
/lab_host="E.coli"
/clone_lib="QG ABCDI lettuce salinas"
/notes="Vector: pBRCDNASFIAB; The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgdb.ucdavis.edu/
TAG_TISSUE="chemical induction"
TAG_LIB="QG ABCDI lettuce salinas"
TAG_SEQ="TGTCAGCCGGG"

ORIGIN

Query Match 21.3%; Score 351; DB 13; Length 692;
Best Local Similarity 69.8%; Pred. No. 1e-77;
Matches 474; Conservative 0; Mismatches 205; Indels 0; Gaps 0;

QY 109 AACAGATAGTGAATGATTAAAGGAGGCGGCACTACTAAAGAGCTTTGGATA 168

14	ATCGAATAGTCCATGATCTCGAAGAGAAAGTGGAGAAAGATATACTGGCTGCTTTATATG	73
169	TTCTTATGAACAATGCGCAATTTGTTGAAGCTGATGATGAAATTCACGCGCTTTGGAATAC	228
74	TTCCAATGGAAACATACAAAACCTCTTGAACCTAATTCGACACAATCCACGCTTTTGGAAATAG	133
229	CGTATCACCTTTGAACGGGAGATTGATCATCGAATTGCGAATGTTATTTATGAACACATATGGTG	288
134	CCTACTATTTTCGACCAAGAGATTAAAGCAAGCCTTGCAGCATATCTATGTGAAATATGGTG	193
289	ATAACTGGAAATGGTGACCGCTCTTTCCTTATGTTTCGGTCTTATTCGAAAGCAAGGATATT	348
194	ATAACTGGAGTGGTGTAGTCTTCTCTGTTTGGTTTTCGACTCATGCGACACAACAGGGTCTT	253
349	ATGTTACATGTGATGTTTTCATAATCATATAAGACAAAAATGGAGCGTTCAAGCAATCGT	408
254	ACGTTTCATGTGATATTTTCCAAATACTACAAAAAAAGAAATGGAGCTTTTAAAGGAATCGT	313
409	TAGCTAATGATGTTGAAGGTTTGCCTTGAGTTGTACGAAGCAACTTCTATCAGGGTACCTG	468
314	TGACCAATGACATCCATGAGATGCTTGAGTTATATGAGGCAACATTTATGAGAGTGAAG	373
469	GGGAGATTATATTAGAAGATGCTCTTGGTTTTACAGATCTCGTCTTTCGATTTATGACAA	528
374	GAGAAGTTGTACTAGACGAAGCTCTTCTTTTCACAAAAACTCATCTTGA AAAA ACTAGCAA	433
529	AAGATGCTTTTCTACAAACCCGCGCTCTTTTACCGAAATACAAACGGGCACTAAAGCAAC	588
434	AGGATCTGTTCGGTGTAATCTACTCTTTTCAATACACATACAGGAGACACTAAAAATC	493
589	CCCTTTGGAAAAAGTTGCCAAGAAATAGAGCGGCGAGTACATTCCTTTCTATCAACAAC	648
494	CCATATTGAAAGGTTGCCAAGACTTGGAGCAATTCACATACATTCCTTTCTACCAAAAAC	553
649	AGAATTCTCATACAGACACTTTACTTAACTTCGTAAAGTTAGAGTTCAATTTGCTTCAGT	708
554	AAGCTTCTTGTAAATGATCTTTACTTCAACTTTCAAAGTTAGGGTTCAATCTACTTCAAT	613
709	CATTGCAACAGGAAGAGCTCAGCCATGTGTGCAAAATGGTGGAAAGCTTTTCGATATCAAGA	768
614	CCCTACACAGAAGAGCTTATGTGAACCTTCCAAAGTGGTGGAAAGGTTTTGATGTACCGA	673
769	AGAACGCACCTTGTTTAAG	787
674	AACAGATACCTTATACAG	692

RESULT 3				
BQ865374				
LOCUS	BQ865374	699 bp	mRNA	linear
DEFINITION	QGC4a04.yg.ab1 QG-ABCDI lettuce salinas	Lactuca sativa	EST 14-AUG-2002	cDNA clone
	QGC4a04.	mrna		sequence.

ACCESSION	B0865374	
VERSION	B0865374.1	GI:22250839
KEYWORDS	Est.	
SOURCE	Lactuca sativa	
ORGANISM	Lactuca sativa	

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Asterales; Asteraceae; Cichorioideae;
Cichorieae; Lactuca.

REFERENCE 1 (bases 1 to 699)

AUTHORS Kozik, A., Michelmore, R.W., Knapp, S., Matvienko, M., Rieseberg, L., Lin, H., van Damme, M., Lavelle, D., Chevillon, P., Ziegler, J., Ellison, P., Kolkman, J., Livingston, M.S., Livingston, K., Zhou, Y., Lai, Z., Church, S., Jackson, L. and Bradford, K.

TITLE Lettuce and Sunflower ESTs from the Compositae Genome Project

**JOURNAL
COMMENT**

THE

QY 709 CATTGCAAGAGAGAGCTCAGCCATGTGTCAAAATGTTGGAAAGCTTTTCGATATCAAGA 768
Db 613 CCTTACCAAGAAAGAGAGCTTAGTGAACCTTCCAGTGGTGAAGAGTTTGTGATGATACCGA 672
QY 769 AGACGACACCTTGTGTAAGATAGA 794
Db 673 ACAAGATACCTTATACACGGGACCGA 698

RESULT 4
BU027611
LOCUS QHG6K24.yg.ab1 OH EFGHJ sunflower RHA280 Helianthus annuus cDNA
DEFINITION QHG6K24, mRNA sequence.
ACCESSION BU027611
VERSION BU027611.1 GI:22463131
KEYWORDS EST
SOURCE Helianthus annuus (common sunflower)
ORGANISM Helianthus annuus

REFERENCE 1 (bases 1 to 697)
AUTHORS Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L., Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J., Ellison,P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z., Church,S., Jackson,L. and Bradford,K.
TITLE Lettuce and Sunflower ESTs from the Compositae Genome Project
JOURNAL Unpublished (2002)
COMMENT Contact: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Amundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@ucdavis.edu [michelmore@vegmail.ucdavis.edu]
singleton, see http://cgpdb.ucdavis.edu/ for details.
Plate: QHG6 row: K column: 24.

FEATURES
source
1..697
/organism="Helianthus annuus"
/mol_type="mRNA"
/cultivar="RHA280"
/db_xref="taxon:4232"
/clone="QHG6K24"
/lab_host="E.coli"
/clone_lib="OH EFGHJ sunflower RHA280"
/note="Vector: pBRCDNASFIAB; The library was constructed from 11 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at http://cgpdb.ucdavis.edu/ TAG_TISSUE=roots environmental stress
TAG_LIB=OH EFGHJ sunflower RHA280
TAG_SEQ=ATCTCGCGGG"

ORIGIN
Query Match 19.7%; Score 325.2; DB 13; Length 697;
Best Local Similarity 67.7%; Pred. No. 3.4e-71;
Matches 472; Conservative 0; Mismatches 219; Indels 6; Gaps 1;

QY 463 TACCTGGGGAGATTATTAGAGATGCTTCTGGTTTACAGATCTCGTCTTAGCATTA 522
Db 1 TGGCAGGCGGAGCTATAGATGATATTTCTGGCTTCACAAAAACCAACTTCACAAA 60
QY 523 TGACAAAGATGCTTTTCTCAAC-----CCGCTCTTTTACCGAAATACACGGG 576

Db 61 TAACAAAGGACCCCTTAGATGGAACCTGTACTCTTCTTTCGATCAGATAGAGGAG 120
QY 577 CACTAAAGCAACCCCTTTGCAAAAAGGTTCGAAGAATAGAGGGCGCGCAGTACATTCCTT 636
Db 121 CACTAGAGCGACCCATATGGAAGAAGATTGCCAAGGCTAGATACGTTGGATACATCCTT 180
QY 637 TCTATCAACAAAGAGATTCTCATACAAAGACTTTTCTTAAACTTGTCTTAAGTTAGATTCA 696
Db 181 TCTATGAACACGAAGATTTCATATGATCGCTAGTAAGACTCGCAAGATTGGATTTC 240
QY 697 ATTTCCTTCAGTCATTCACAGGAGAGCTCAGCCATGTGTGCAAAATGTTGGAAAGCTT 756
Db 241 ACAGGCTTCAGTCATTCACAAAAAGAGCTTAGCCAACTTGCACAGTGGTGGAAAGCTT 300
QY 757 TCGATATCAAGAAGAACGACCTTGTGTTTAAAGAGATAGAAATGTTGAATGCTTACTTTGGG 816
Db 301 TTGACCCCTCCAAAGATCTGCATATCTAAGAGACAGAAATGTTGAAGGATATTTCTGGA 360
QY 817 GACTAGGTTCAAGCTATGAGCCACAGATTATTCCTGGGCTAGAGTTTCTTCAAAAAGCTG 876
Db 361 TAGTGGGTGCTACTTTGAGCCCTCAATATTTCTCGGTCAGAAATTTCTTACAAAAATGT 420
QY 877 TTGCTCTTAACTCTTATAGATGACACTTATGATGCTATGCTATGCTTATGAAAGACTTA 936
Db 421 TTAATAATGTTATCAATTTTAGATGACACTTATGATCTTATGGTCTTTATGAGACTTG 480
QY 937 AGATCTTTTACTGAAGCTGTTTGAAGGTGGTCAATTATCATGCTTAGACACACTTCCAGAA 996
Db 481 AGATTTTCCACCAAGCAGTTCAAAGCTGGTCAATCGCTGCATGATGCTTCTCTGATT 540
QY 997 ACATGAACCGATATACAAATTTATTCATGATACATACACAGAAATGGAAGAAATTTCTTG 1056
Db 541 ACATGAAGCTGTCTATACAAAATACTCTTGGATTGTTTACCATGAATAGAGGAAATATG 600
QY 1057 CAAAGGAGGGAACACAGATCTATTTAACTCGCGCAAGAAATTTGTGAAAGAGTTTGTTA 1116
Db 601 CAAACCGAAGAAAGCATATCAAGTTCCATGCCCGAGAGTTGNATCAAGAGATGAGTA 660
QY 1117 GAAACCTGATGTTGAAGCAAAATGGCAAAATGAGGG 1153
Db 661 GATGCTACATGATTGAAGCAAAATGGAGAAACGACGG 697

RESULT 5
CF211234
LOCUS CF211234
DEFINITION CF211234, 814 bp mRNA linear EST 01-AUG-2003
ACCESSION CAB20007 Iva Fa G05 Cabernet Sauvignon Flower bloom - CAB2 Vitis
VERSION vinifera cDNA clone CAB20007_Iva_Fa_G05 5', mRNA sequence.
KEYWORDS CF211234.1 GI:33405607
SOURCE Vitis vinifera
ORGANISM Vitis vinifera

REFERENCE 1 (bases 1 to 814)
AUTHORS Goes da Silva,P., Iandolino,A., Lim,H., Baek,J., Jones,K. and Cook,D.
TITLE Expressed sequence tags from Vitis vinifera 'Cabernet sauvignon' berries at various developmental stages
JOURNAL Unpublished (2003)
COMMENT Contact: Douglas Cook, PhD
CAES Genome Facility
UC Davis, Plant Pathology
One Shields Ave, Davis, CA 95616, USA
Tel: 530 754 6561
Fax: 530 754 6617
Email: drcook@ucdavis.edu
Seq primer: ACGTACCGGACATATGCC.
Location/Qualifiers
1..814
/organism="Vitis vinifera"

/mol_type="mRNA"
/cultivar="Cabernet Sauvignon"
/db_xref="taxon:29760"
/clone="CAB20007 IVa_Fa_G05"
/sex="Hermaphrodite"
/dev_stage="Bloom"
/clone_lib="Cabernet Sauvignon Flower bloom - CAB2"
/note="Organ: Flower - Bloom; Vector: pDNR; Site: 1; Sfil;
Site 2: Sfil; CAB2 is a cDNA library of Vitis vinifera cv.
'Cabernet Sauvignon' Clone 8 berries. Samples were
collected at full bloom (80 to 100% flowers showing
dehiscence of calyptas or caps and anthers fully
extended). Samples were located at the University of
California, Davis, Experimental Vineyard. cDNAs were made
by oligo-dT priming and directionally cloned. 5' and 3'
adaptors were used in cloning as follows:
5'-AAGCAGTGGTATCAACGAGGCGGCGCATTCAGCGCGG-3' and
5'-ATTCTAGAGCGGCGGCGGCGCATG-dt(30)NN-3'. Library was
constructed using the Clontech Creator SMART kit and
size-selected to contain the 0.5-3 kb size fraction."

1496 AAACATTCCAGGCGGTATTGATGGCTGTGATCTATTGTGCGAGTTCTTGAAGTTCA 1555
725 CGCTGTTTCAATGCCCTCTCTTGTGCTGATTTCAATCTTTCAGAGCCGCGGATGTCAT 784
1556 ATATCGAGGAGAGGATAACTTCACACGTAT 1585
785 TTACAAAGAGATGATTCGTACACATGT 814

RESULT 6
LOCUS CB970399/c 742 bp mRNA linear EST 30-APR-2003
DEFINITION CAB10003_IVa_Ra_F11 Cabernet Sauvignon Flower Pre-bloom - CAB1
Vitis vinifera cDNA clone CAB10003_IVa_Ra_F11 3', mRNA sequence.
ACCESSION CB970399
VERSION CB970399.1 GI:30252848
KEYWORDS EST
SOURCE Vitis vinifera
ORGANISM Vitis vinifera
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; Vitaceae; Vitis.
1 (bases 1 to 742)
Goes da Silva,F., Iandolo,A., Lim,H., Baek,J., Jones,K. and
Cook,D.
Expressed sequence tags from Vitis vinifera 'Cabernet Sauvignon'
berries at various developmental stages
Unpublished (2003)
Contact: Douglas Cook, PhD
CAES Genome Facility
UC Davis, Plant Pathology
One Shields Ave, Davis, CA 95616, USA
Tel: 530 754 6561
Fax: 530 754 6617
Email: drcook@ucdavis.edu
Seq primer: GCCAAGCAATGGCTTAG.
Location/Qualifiers
1..742
/organism="Vitis vinifera"
/mol_type="mRNA"
/cultivar="Cabernet Sauvignon"
/db_xref="taxon:29760"
/clone="CAB10003 IVa_Ra_F11"
/sex="Hermaphrodite"
/dev_stage="Pre-bloom"
/lab_host="DH5alpha"
/clone_lib="Cabernet Sauvignon Flower Pre-bloom - CAB1"
/note="Organ: Flower - Pre-bloom; Vector: pDNR; Site: 1;
Sfil; Site 2: Sfil; CAB1 is a cDNA library of Vitis
vinifera cv. 'Cabernet Sauvignon' Clone 8 flowers. Samples
were collected approximately eleven days before onset of
bloom (clusters at this stage were fully developed and
flowers with calyptas or caps still attached. Sampled
vines were located at the University of California, Davis,
Experimental Vineyard. cDNAs were made by oligo-dT priming
and directionally cloned. 5' and 3' adaptors were used in
cloning as follows:
5'-AAGCAGTGGTATCAACGAGGCGGCGCATTCAGCGCGG-3' and
5'-ATTCTAGAGCGGCGGCGGCGCATG-dt(30)NN-3'. Library was
constructed using the Clontech Creator SMART kit and
size-selected to contain the 0.5-3 kb size fraction."

Query Match 15.6%; Score 256.6; DB 14; Length 742;
Best Local Similarity 60.6%; Pred. No. 7.5e-54;
Matches 443; Conservative 0; Mismatches 279; Indels 9; Gaps 1;
319 GGTTCGCTCTTATCGAAGCAAGGATATTATTGTTACATGATGATTTTCAATACTATA 378
740 GATTTGACTACTAGGCAACAGGGTACACTATTTCATGCGATATATTCACAAAGTTTA 681
379 AAGACAAAATGAGCGCTTCAGCAATCGTTAGCTATGATGTTCAAGGTTTGTGAGT 438

/mol_type="mRNA"
/cultivar="Cabernet Sauvignon"
/db_xref="taxon:29760"
/clone="CAB20007 IVa_Fa_G05"
/sex="Hermaphrodite"
/dev_stage="Bloom"
/clone_lib="Cabernet Sauvignon Flower bloom - CAB2"
/note="Organ: Flower - Bloom; Vector: pDNR; Site: 1; Sfil;
Site 2: Sfil; CAB2 is a cDNA library of Vitis vinifera cv.
'Cabernet Sauvignon' Clone 8 berries. Samples were
collected at full bloom (80 to 100% flowers showing
dehiscence of calyptas or caps and anthers fully
extended). Samples were located at the University of
California, Davis, Experimental Vineyard. cDNAs were made
by oligo-dT priming and directionally cloned. 5' and 3'
adaptors were used in cloning as follows:
5'-AAGCAGTGGTATCAACGAGGCGGCGCATTCAGCGCGG-3' and
5'-ATTCTAGAGCGGCGGCGGCGCATG-dt(30)NN-3'. Library was
constructed using the Clontech Creator SMART kit and
size-selected to contain the 0.5-3 kb size fraction."

Query Match 16.3%; Score 269.2; DB 14; Length 814;
Best Local Similarity 58.3%; Pred. No. 5e-57;
Matches 472; Conservative 0; Mismatches 338; Indels 0; Gaps 0;
776 ACCTTGTTTGAAGATAGAATTGTTGAATGCTACTTTTGGGACTAGGTTTCAGGCTATGA 835
5 ACCTTTGCAGGACAGATTGGTTGAATGCTACTTTTGGGATGCTTGGGCTGATTTTGA 64
836 GCACAGATATCCCGGCTAGAGTTTCTTCACAAAGCTGTTGCTGTATTAATCTTAT 895
65 GCCTCAATACTTACGGGCTAGACGAATCTTAACCAAGTAATGCCATGACATCCATCT 124
896 AGATGACATTTATGCTGCTATGCTATGATGAAGACTTAAGTCTTTTACTGAAGCTGT 955
125 AGATGATATCCATGATGATGACACACCTGAAGACTCAAGCTCTTCATAGAGCCAT 184
956 TGAAGGTGTTCAATCTTAGACACACTTCCAGAAATACATGAACCGATATCAAA 1015
185 TGAGAGTGGGATATTGACAGTATAGATCAGCTCCAGAAATACATGAATCTGCTATT 244
1016 ATTATTCATGATATACATACAGAAATGGAAGATTTCTTGAAGGAGGAGGAAGACAGA 1075
245 AGCACTCTTAGATGTGTACAAAGAAATCGAGAGAGATGGAAGAAAGGAAACCAATA 304
1076 TCTATTTAATCGCGCAAGATTTGTGAAGAGTTTGTGAACCTGATGTTGTAAGC 1135
305 TCGGTTTCACTATGCCAAAGAAATGAAGAAATCAAGTTCGAGCTTACTTTGCTGAGGC 364
1136 AAAATGGCAAAATGAGGACACATACCAACCACTGAAGAGCATGATCCAGTTGTAATCAT 1195
365 TAATGTTTACATGAGAACACGTACCAACATTTGAAGAGTACATGCGTGTGCACTAGT 424
1196 TACTGGGCGTCTAACCTGCTTACAACTTTGTTATCTTGGCATGAGTGATATTTCAC 1255
425 AAGCTCTGGTACTGCTATGCTGCAACACGCTCTTCTGTTGGAAATGGGAAATAGCAAC 484
1256 AAAAGATCTGTCGAATGGGCTGCTCTGCACTCTCTTTTATAGTACTCAGGTATATCT 1315
485 AAAGGAGCGCTTTGATTTGGGTGACAGCAATCCTAAGATTTTGTCTTCAAAATTTAT 544
1316 TGGTCGAGCGCTCAATATGATCTCATGACCCCAAGGCGGAGCAAGAAACAAATAGTTTC 1375
545 TATAAGGCTCATAGTATGATCAATCAATCATTAAGTTTGAGCAAGAGGCGCATGTGC 604
1376 ATCAGAGCTTGAAGTTTATGAAGAAATATATGTCATAGGAGTATGCCCAACCTT 1435
605 CTCAGCTTTGAATGTTATCAATGAAGCAATATGTTGCTCTGAGGAAACAGTATACAGTGA 664
1436 GATTTCAGGAGTATGAGATGTTGGAAGCAATATAACCGAGTACTCACAACCTAA 1495
665 ATTTCAAAAGCAAAATTTGAGATGCTAGTGTGCTGATATTAACAGGAATGCTTCAACCTAC 724

Db 680 CGGATGAACGAGGTAGATTCAAGGAAGCTTTGATCAGCGATGTAAGAGGCGATGCTAGGCT 621
Qy 439 TGTACGAGCAACTTCTATGAGGGTACCTGGGAGATTATATAGAGATGCTCTTGGTT 498
Db 620 TGTATGAAGCTGACATCTGAGGGTTCATGGAGAGACATCTTGCAGAGCACTTGTCT 561
Qy 499 TTACACGATCTCGTCTTACGATTAATGACAAAGAGTCTTTTCTACAAACCCCGCTCTTT 558
Db 560 TCACACCACTCATCTCAAGGCCATGTTAGAAAGTTAGGATAT-----CATCTTG 510
Qy 559 TTACGGAATACAGCGGCACTAAAGCAACCCCTTTGGAAAGGTTGCCAAGATAGAGG 618
Db 509 CAGAACTAGTTGCTCATGCCCTGAACCGGCCCATAGAAAGGTTGGAGAGCTAGAGG 450
Qy 619 CGGCGCAGTACATCTCTTCTATCAACACAGATTTCTATCAACAGACTTTTACTTTAAAC 678
Db 449 CAAGATGTTATATCTGTCTACCAAGATGAAGCTTTCCATGATATAAATTTTACTAGAGC 390
Qy 679 TTGCTAAGTTAGAGTTCAATTTGCTTCACTCATGTCACAAAGAGAGCTCAGCCATGTGT 738
Db 389 TAGCAAAATTAGATTTCAATCTAGTGCAGTCACTGCACAAAGAGAGCTAAGCAATCTTG 330
Qy 739 GCAATGTTGAAAGCTTTTCGATATCAAGAAAGACGCACTTGTTTTAAAGATAGAAATG 798
Db 329 CAAGGTGTTGAAAGATTTAGACTTTTGTACAAAGTTTACCTTTTGCACGAGACAGATTGG 270
Qy 799 TTGAATGCTACTTTTGGGAGCTAGGTTCAAGGCTATGAGCCACAGATTCCCGGGCTAGAG 858
Db 269 TTGAAGGCTACTTTTGGATGTCATGAGGCTATTTTGGAGCCCAATCTTACGGGGTAGAC 210
Qy 859 TTTTCTTCAAAAAGCTGTGTGTTATTAACCTTTATAGATGACACTTTATGATGCGTATG 918
Db 209 GAATTTCAACCAAGTAATGTCATGACATCCATTTTATGATATCCATGATGATATG 150
Qy 919 GTACTTATGAAGAACTTAAAGTCTTTTCTGAAAGCTGTTGAAAGGTTGCTCAATTACATGCT 978
Db 149 GCACCCCTGAAGAACTCAAGCTCTTCTATGAAGACCAATTTGAGAGATGGGATATTAACAGCA 90
Qy 979 TAGACACACTTCCAGAAATACATGAACCGATATACAAATTTTATGATATACATACACAG 1038
Db 89 TAAATCAGCTTCCAGAAATACATGAAGAACTCTGCTATGTCGCACTTTAGATGTGTACAAG 30
Qy 1039 AAATGGAAGAA 1049
Db 29 AAAAAAAAAA 19

RESULT 7
CB913201
LOCUS
DEFINITION
CB913201 634 bp mRNA linear EST 25-APR-2003
VLD163B02 376459 An expressed sequence tag database for abiotic stressed berries of Vitis vinifera var. Chardonnay
CDNA clone VLD163B02 5, mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

CB913201
EST.
Vitis vinifera
Vitis vinifera
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; Vitaceae; Vitis.
1 (bases 1 to 634)
Cushman, J.C.

REFERENCE
AUTHORS
TITLE
An expressed sequence tag database for abiotic stressed berries of Vitis vinifera var. Chardonnay

JOURNAL
COMMENT

Unpublished (2002)
Contact: Cushman JC
Department of Biochemistry
University of Nevada
MS200, Reno, NV 89557-0014, USA
Tel: 775-784-1918
Fax: 775-784-1650

Email: jcushman@unr.edu

PCR Primers
FORWARD: T3 20mer
BACKWARD: T7 21mer (backward)
Plate: 163 row: B column: 02
Seq primer: T3 20mer
High quality sequence stop: 634.
Location/Qualifiers

FEATURES
source

1..634
/organism="Vitis vinifera"
/mol_type="mRNA"
/db_xref="taxon:29760"
/clone="VVD163B02"
/tissue_type="berries"
/dev_stage="mixed; 8, 9, 11, 13, 15, 16 weeks daf"
/clone_lib="An expressed sequence tag database for abiotic stressed berries of Vitis vinifera var. Chardonnay"
/note="Vector: lambda Uni-Zap XR, Bluescript SK-; Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Query Match 15.0%; Score 246.6; DB 14; Length 634;
Best Local Similarity 63.0%; Pred. No. 2.4e-51;
Matches 381; Conservative 0; Mismatches 224; Indels 0; Gaps 0;
Qy 588 CCCCTTTGAAAAAGTTGCCAAGAAATAGAGCGCGCGAGTACATTCCTTTCTATCAACAA 647
Db 6 CCCATTAGAAAAAGTTTGGAGAGGCTAGAGGCAAGATGGTATATCTGTCTACCAAGAT 65
Qy 648 CAGATTTCTCATACAAAGACTTTTAACTTCTAGTTAGAGTTCAATTTGCTTCTAG 707
Db 66 GAAGCTTTCCATGATATAAAGCTTTTACTAGAGCTAGAGAAATAGATTTCAATCTAGTGCAG 125
Qy 708 TCATTGCACAAAGAGAGCTCAGCCATGTGTGCAAAATGGTGGAAAGCTTTTCGATATCAAG 767
Db 126 TCATGCAACAAAGAGAGCTTAGCAATCTTGCAGAGGTGGTGGAAAGATTAGACTTTGCT 185
Qy 768 ARGAAACCACTTGTGTTTAAAGAGATAGAAATTTGAAATGCTACTTTTGGGGAGTACGTTCA 827
Db 186 ACAAGTTACTTTTGCACGAGACAGATTGGTTCAAGGCTACTTTTGGATGCAATGGGGTG 245
Qy 828 GGCATGAGCAGCAGTATTCGCGGCTAGAGTTTCTTCAAAAGAGCTTTGCTGTATATA 887
Db 246 TATTTTGGAGCCCAATACTTACGGGGTAGAGCAATTTCAACAAAGTAAATGTCATGACA 305
Qy 888 ACTCTTATAGATGACACTTATGATGCGTATGCTTATGAAAGACTTAAAGATCTTTTACT 947
Db 306 TCCATTTCTAGATGATATCCATGATGATGCGCACACCTGGAAGACTCAGGCTCTTCATA 365
Qy 948 GAAGCTTTGAAAGGTGGTCAATTAATGCTTTAGACACACTTCCAGATACATGAACCG 1007
Db 366 GAAGCATTGAGAGATGGGATATTAAACAGCATAAATCAGCTTCCAGAAATACATGAACCTC 425
Qy 1008 ATATACAAATTTTATGATGATACATACACAGAAATGGAGAAATTTTCTTCAAGAGGAGGA 1067
Db 426 TGCTATGTCGCACTCTTAGATGTGTACAAAGAAATCGAGGAGAGATGGAGAAAGAGGA 485
Qy 1068 AGAAGATCTTATTTAACTGCGCAAGAAATTTGTGAAGAGTTTGTAGAAACCTGATG 1127
Db 486 AACCAATTCGGTTTCACTATGCCAAGAGATTAATGAAGAAATCAAGTTCCGAGCTTACTTT 545
Qy 1128 GTTGAACCAAAATCGGCAATAGGGGACATACCAACCACTGAAGAGATGATCCAGTT 1187
Db 546 GCCGAGGCCAAATGTTTACATGAAGAAACACGTCAGCAATTTGAAGAGTACATGCGGTGT 605
Qy 1188 GTAAT 1192
Db 606 GCACT 610

RESULT 8
CD486420
LOCUS

779 bp mRNA linear EST 01-JUL-2003

CD486420

```

DEFINITION CRH1.4G10 Cotton Root and Hypocotyl Lambda ZIPLOX Library (CRH)
Gossypium hirsutum cDNA clone CRH1.4G10 5', similar to
(+) -delta-cadinene synthase, mRNA sequence.
ACCESSION CD486420
VERSION CD486420.1 GI:31407395
KEYWORDS EST.
SOURCE Gossypium hirsutum (upland cotton)
ORGANISM Gossypium hirsutum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eucosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
1 (bases 1 to 779)
Dowd,C., Wilson,I. and McFadden,H.
Different Gene Expression Responses in Cotton Root and Hypocotyl
tissues during infection with Fusarium Wilt Disease
Unpublished (2003)
Contact: Caitriona Dowd, Helen McFadden
Commonwealth Scientific and Industrial Research Organisation
Division of Plant Industry
Black Mountain Laboratories, Cnr Clunies Ross Street & Barry Drive,
Black Mountain, Canberra, ACT, 2601, Australia
Tel: 61 2 6246 4914, 6246 5377
Fax: 61 2 6246 5000
Email: Caitriona.Dowd@csiro.au, Helen.McFadden@csiro.au
Vector clipped sequences Bases 1-17 (GTCGACCCACGGCTCG): SalI
adapter
Seq primer: M13 reverse primer
High quality sequence stop: 779.
Location/Qualifiers
1. 779
/organism="Gossypium hirsutum"
/mol_type="mRNA"
/cultivar="DeltaEMERALD"
/db_xref="taxon:3635"
/clone="CRH1.4G10"
/tissue_type="Root and hypocotyl tissues"
/dev_stage="5 day old seedlings"
/lab_host="Y1090(ZL)"
/clone_lib="Cotton Root and Hypocotyl Lambda ZIPLOX
Library (CRH)"
/notes="Vector: Lambda ZIPLOX; Site 1: SalI; Site 2: NotI;
mRNA was prepared from root and hypocotyl tissues of the
cotton cultivar DeltaEMERALD. cDNA was synthesised from a
NotI-oligo primer/adaptor using the manufacturers
protocols (Life Technologies) and then ligated to a SalI
adaptor to facilitate directional cloning. The cDNA was
cloned into the SalI and NotI sites of the Lambda ZIPLOX
phage vector (Life Technologies). Constructed by Caitriona
Dowd and Helen McFadden."

ORIGIN
Query Match 14.8%; Score 244; DB 14; Length 779;
Best Local Similarity 59.3%; Pred. No. 1.2e-50;
Matches 456; Conservative 0; Mismatches 301; Indels 12; Gaps 2;

QY 228 CCGTATCATTGACGGGAGATGATCATGTCATGTAATGTAATTAAGAACAATATGCT 287
DB 19 CCGGTTCATTTCACGAGGATGCGAGATGATGATGATATCTCATTACACACAT 78
QY 288 GATAACTGGAATG---GTGACCGCTCTTCCTTATGTTTCCGTCTTATGCGAAAGCAAGGA 344
DB 79 GATGCGGAGAACGACCTCTACACCACATCCCTTCGATTCGACTACTCCGAGAGCATGGA 138
QY 345 TATTATGTTACATGATGATGTTTCAATAACTATAAAGCAAAATGAGCGCTTCAGCA 404
DB 139 TTCATATGTTTCATGCGACGATTTCAACAAAGTTTAAAGACGAGGGAATTTCAAGTCA 198
QY 405 TCCTTAGCTAATGATGTTGAAGGTTTGTCTGAGTTGTACGAGCAACTTCTATGAGGGA 464
DB 199 TCGTGTACAGCGAATGTTGAGGATGTTGGAATTTACCAAGCTTCTATTTGAGGTT 258
QY 465 CCTGGGAGATTTATTAGAGATGCTCTTGTTGTTTACACCATCTCGTCTTAGCATTTATG 524
```

```

Db 259 CATGGGGAAGATATATTGGATGAAGCAATTTCTTTCACAGCAACCAATTTAAAGCCTT--- 315
QY 525 ACAAAAGATGCTTTTCTACAAACCCCGCTCTTTTACCGAAATACAAACGGGCACTAAAG 584
Db 316 -----GCAGTAGCATCTTTGGACCATCTTTATCCGAAGAGGTTTCTCATGCTTTGAAA 369
QY 585 CAACCCCTTTGGAAAGGTTGCAAGATAGAGGGCGCGAGTACATTCCTTTCTATCAA 644
Db 370 CAATCAATTCGAAGAGGCTTGCCAAAGGTTGAGGCAAGACACTATCTTTTCAGTATACCAA 429
QY 645 CAACAGATTTCTCATACCAAGACTTTTACTTAAGTTGCTAAAGTTAGAGTTCAATTTGCTT 704
Db 430 GATATTGAGTCCCATAAATAGGTTTGTGGAGTTTGTCTAAGATCGATTTCAACATGGTA 489
QY 705 CAGTCATTGCAACAAGGAGAGCTCAGCCATGTGTGCAAAATGTGTGAAAGCTTTTCGATATC 764
Db 490 CAACCTTTTGCATAGAAAAGAGCTAAGTGAGATTTCTAGGTGTGTGGAAGGATTTAGACTTT 549
QY 765 AAGAAGACGCACCTTTGTTTAAGAGATAGAAATTTGTAATGCTACTTTTGGGGAAGT 824
Db 550 CAAGAAAGTTGCCATATACGCAAGAGATAGAGTGGTTGAAAGGCTATTTTTCGATCTCAGGA 609
QY 825 TCAGGCTATGAGCCACAGATTTCCCGGGCTAGAGTTTCTTCACAAAGCTGTGCTGTT 884
Db 610 GTGTACTTTGAGCCCCCAATATTTCTTGTGTAGAAAGATGTTGACAAAGTATAGCAATG 669
QY 885 ATAACTCTTATAGATGACACTTATGATGCTATGCTATGCTATGATGCTATGATGCTATGCT 944
Db 670 GCATCTATTGTAGATGATACATATGACTCATATGCAACATATGAGAGCTCATTCCTCTAT 729
QY 945 ACTGAAGCTGTTGAAAGGTGGTCAATTAATGATGCTTAGACACACTTCCAG 993
Db 730 ACAATGCAATTTAGAGGTGGGATATCAAAATGATAGTGAATCTTCTGCT 778

RESULT 9
AY105371
LOCUS AY105371 1981 bp mRNA linear HTC 16-OCT-2002
DEFINITION Zea mays PCO100777 mRNA sequence.
ACCESSION AY105371
VERSION AY105371.1 GI:21208449
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
Ratney,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
Unpublished (2002)
2 (bases 1 to 1981)
Direct Submission
Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
If you are interested in getting corresponding physical clones,
these are publicly available from ZmDB, www.zmdb.iastate.edu; TIGR,
searching at MSL, maizegap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
maize cDNA sequences is either Virginia Walbot, Stanford or Pat
Schnable, Iowa State, then clones may be requested from ZmDB:
www.zmdb.iastate.edu.
location/Qualifiers
1. 1981
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:635786"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Cornsensus
Library"

FEATURES
source
```

/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"

ORIGIN

Query Match 14.7%; Score 243.2; DB 11; Length 1981;
Best Local Similarity 50.5%; Pred. No. 2.3e-50;
Matches 739; Conservative 0; Mismatches 698; Indels 27; Gaps 5;

QY 38 TGCCAACTTCTCCCAAGCAATTTGGGGAGATGAGTTTCTCTATCTATCAAAAGCAAGTAGA 97
DB 113 TCCTCTCTCCATCCACTCTGTGGGAGATTTCTTTCTTTTACGAGGCGCCAGCGA 172
QY 98 GCAAGGGGTGGAACAGATAGTGAATGATTTAAAGAAAGTGGCGCACTACTAAAGA 157
DB 173 G---GCACAGGAAGCTGAAATGCGCAAGGGCTGAAGTGTGGGGAAGAGTGAGAAA 229
QY 158 AGCTTTGGATATTCCTATGAACATGCGCAATTTGTTGAAGCTGATTGATGAAATTCACG 217
DB 230 CATGATTAAGGTTCAATGATGTGCCAGAAATAGTGACCTTATATCACCTGCAGCG 289
QY 218 CTTTGAATACCGPATCACTTTTGAACGGGAGATTGATCATGCAATGCAATGATTTATGA 277
DB 290 ACTAAATCTGACTACCACTACGAGGATGAGATAAATGAAAAGTTGGCTGTGTTTACAA 349
QY 278 AACATATGGTGATTAACGGAATGTTGACCGCTCTTCCTATGTTGGTTCGCTTATCGGAA 337
DB 350 CTCGAATATGATGGTGTAAATTTGGACTTATGTTTTCACGCGGATTTCTATCTTCGCTAA 409
QY 338 GCAAGGATATTTATGTTACATGTGATGTTTCAATAACTATGAACAAAGATGGAGCGTT 397
DB 410 ATGTGGCTATCATGTTTCATCGATGATTTCTTAATTTCAAGACCAATATGGAATTT 469
QY 398 CAAGCAATCGTTAGTAAATGATTTGAAGTTTGGCTTGAAGTTGATGATGACGAACCTTAT 457
DB 470 TA-----TTGAGGTAGATACCGAGGCTTGTAAAGCTTATACAATGCGACATACCT 520
QY 458 GAGGTACTCTGGGAGATATATTAGAGATGCTCTTGGTTTACAGATCTCTGCTTAG 517
DB 521 GAGGATTCATGGAGACAGCTGCTGTGATGAGCAATTTCTTTTACGAGATGCTTCA 580
QY 518 CATATGACAAAAGATGCTTTTCTACAAACCCGCTCTTTTATCCGAAATCAACGGGC 577
DB 581 AGATAGATTAGAACATTTGGATCACCAATAGCAG-----AGGAAGTCTGCTGTC 631
QY 578 ACTAAGCAACCCCTTTGGAAAAGTTGCCAAGAAATAGAGCGCGCGCAAGTACATTCCTTT 637
DB 632 CTTTGATACCCCACTTTTGAAGGTTTGAAGCTTTGGAATGGAAGATTACATTCCTAT 691
QY 638 CTATCAACAACAGATTTCTCATACAAGCTTTACTTAACTTGTCTAAGTTAGAGTTCAA 697
DB 692 CTACGAAAGGACGCTAAACAAACAAATCAATNTTAGAGTTTTCGAAACTGAATCTCAA 751
QY 698 TTGCTTCAGTCATTGCAACGAAGAGCTGACGCAATGTGCAATGTTGGAAGCTTT 757
DB 752 CCTTCTACAGCTTCTTATTCATCAGAGCTCAAGAGTGCAACCTGTTGGAAGAACT 811
QY 758 CGATATCAAGAAAGCAACGCTTTGTTTGAAGATAGAAATGTTGAATGCTACTTTTGGGG 817
DB 812 CCGCTCGAATCAAACTTGAGTTTGTGAGGATAGATAGTAGAGGTGCTACTTTGGAT 871
QY 818 ACTAGGTTCAAGCTATGAGCCACAGTATTCGGGCTAGAGTTTCTTCAAAAGCTGT 877
DB 872 GAGTGGGGATGCTATGATCTCTCAATTTCTCATTTCTCGAATATATCTTACAAAGATTGT 931
QY 878 TGCTGTTTAACTCTTATAGATGACATATGATGCTGATGCTATGATGAAGAACTTAA 937
DB 932 TGCCTTCATCAGATCTAGATGATACCTTCACTCAGATGCTAATTTCTTATGAGGAT 991
QY 938 GATCTTTACTGAAGCTGTGAAAGGTGCTCAATTATCATGCTTAGACACACTTCCAGATA 997

DB 992 GCAACTTGCAGAGCAGTGCAGAGGTGGGATGAAAGTGCAGTCTCTACTCTCCAGATA 1051
QY 998 CATGAAACCCGATATACAAATTTATTCATGGATACATACACAGAAATGGAAGAAATTTCTTGC 1057
DB 1052 CATGAAGGATTTCTACATGATTTTGTGAAGACATTTTCATGCTTTGAGATGATGAGG 1111
QY 1058 AAAGGAGGGAAGAACAGATCTATTTAACTGCGGCAAGAAATTTGTGAAAGATTTGTTAG 1117
DB 1112 ACCAGA---CAAGAGCTACCGAGTGTCTACTTTAAAGAGCGTGAACGGAATTAGTTTCG 1168
QY 1118 AACCTGATGTTTGAAGCAAAATGGCAATGAGGACACATACC---AACCACTGAAGA 1174
DB 1169 GGAATATACCAAGGAATAAATGGCGTGAATGAGGATTTATGTCCTCCAAACATTTGAAGA 1228
QY 1175 GCATGATCCAGTTGTAATCACTTACTGCGGCTGCTAACTCTCTTACACAACTTTGTTATCT 1234
DB 1229 ACACATAAAGTTTCACTCATAGCATTTGAGGTACACTGCTGCTGTTCCGCGTTTGT 1288
QY 1235 TGGCATGAGTATATATTCACAAAAGAGTCTGCGAATGGGCTGTCTCTGCACTCTCTCT 1294
DB 1289 TGGAAATGGGTGATGTAGTAACCAAGAGATCATGGAATGGGTTATGATGATGATGCTGAACT 1348
QY 1295 TTTTATGATCTCAGGTATATCTTGGTGCAGCGCTTAATGATCTCATGCCCAACAGCGCGA 1354
DB 1349 TGTTAAGTCTTTTGGTATATTTGTACGGCTCTCCAATGATTTGTATCAACCAAGCGTGA 1408
QY 1355 GCAAGAAAGAAACATAGTTTCATCGAGCGCTTGAAGTTATATGAAGAAATATAATGTCAA 1414
DB 1409 ACAAGAGAGAGCATTTGTCTCCACTGTTTCACTGTTTACATGAAGCAGATGAACAAAC 1468
QY 1415 TGAGGATGATGCCCAACCTTTGATTTTCAAGGAAGTAGAAGATGTTGGAAGAAATATAAA 1474
DB 1469 AATGGATGAAGCATGTGAGCAGATAAAGAACTTACAGAAAGATTCATGGAAGTTTCATGAT 1528
QY 1475 CCGAGATGATCCACAACTAAATAA 1498
DB 1529 TGAACAAGGCTTGCACATAAAGA 1552

RESULT 10

AU293055 628 bp mRNA linear EST 04-DEC-2002
AU293055 zinnia cultured mesophyll cell equalized cDNA zinnia
LOCUS
DEFINITION
elegant cDNA clone Z7794, mRNA sequence.

ACCESSION AU293055

VERSION AU293055.1

KEYWORDS GI:24253563

SOURCE EST.

ORGANISM Zinnia elegans

Demura, T., Tashiro, G., Horiguchi, G., Kishimoto, N., Kubo, M.,

Matsushita, N., Minami, A., Nagata-Hiwatashi, M., Nakamura, K.,

Okamura, Y., Sassa, N., Suzuki, S., Yazaki, J., Kikuchi, S., and

Fukuda, H.

1 (bases 1 to 628)

Demura, T., Tashiro, G., Horiguchi, G., Kishimoto, N., Kubo, M.,

Matsushita, N., Minami, A., Nagata-Hiwatashi, M., Nakamura, K.,

Okamura, Y., Sassa, N., Suzuki, S., Yazaki, J., Kikuchi, S., and

Fukuda, H.

Visualization by comprehensive microarray analysis of gene

expression programs during transdifferentiation of mesophyll cells

into xylem cells

Proc. Natl. Acad. Sci. U.S.A. 99 (24), 15794-15799 (2002)

Contact: Taku Demura

Morphogenesis Research Group

RIKEN Plant Science Center

1-7-22 Suehirocho, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9605

Fax: 81-45-503-9573

Email: demura@postman.riken.go.jp

This clone was obtained at our laboratory.

Seq primer: M13 forward

Location/Qualifiers

FEATURES

source 1. 628
 /organism="Zinnia elegans"
 /mol_type="mRNA"
 /cultivar="Canary bird"
 /db_xref="taxon:34245"
 /clone="Z7794"
 /tissue_type="mesophyll cell"
 /clone_lib="zinnia cultured mesophyll cell equalized cDNA"
 /notes="Vector: pGEM-T easy; cultured in tracheary element differentiation-inductive medium"

ORIGIN

Query Match 14.5%; Score 239.8; DB 9; Length 628;
 Best Local Similarity 64.1%; Pred. No. 1.3e-49;
 Matches 361; Conservative 0; Mismatches 202; Indels 0; Gaps 0;

591 CTTTGGAAGAGTTGCCAAGATAGAGGGCGCAGTACATCTCTTCTATCAACACAA 650
 Db 1 CTTGGAAGAGTTGCCAAGCTAGAGCGTTGCGTTACATGCAATCTACCAACAA 60
 QY 651 GATTCTCATACAGAGCTTTTACTTAACTGTCTAAGTTAGAGTTCAATTTGCTTCAGTCA 710
 Db 61 GAATCTCATATCATGCTTGTCTAACACATGCTTAAAGTTAGATTTCATCTCCAGAT 120
 QY 711 TTGCACAGAGAGCTCAGCCATGTGTGCAATGTGGAAGCTTTCGATATCAAGAAG 770
 Db 121 TTGCACCGAAGAGCTTTAGCCGAATTTAGCAAGGGTGAAGAAGCTTGTATGCTCAAA 180
 QY 771 AACGCACCTGTTTAAAGATAGATGTTGATGCTACTTTTGGGACTAGCTTCAGGC 830
 Db 181 AAATCTACCTTATGTTTCGGGACAGATGTCGAAGGTTACTTTTGGATATTGGCGGTCTAC 240
 QY 831 TATGAGCCACAGTATTCCTCGGGCTAGAGTTTCTTCCAAAAGCTGTGCTGTATAACT 890
 Db 241 TTGAGGCTCAACATTCGGATCAAGGATTTTCTTATGAAGCTTGCAATCTAGTGATC 300
 QY 891 CTTATAGATGACACTTATGATGCGTATGATGTTATGAGAACTTAAGATCTTTACTGAA 950
 Db 301 ATTTTGATGACATATGATTAATCTATGTTGATGTTATGAGAACTTCGAGATCTTTACAGAA 360
 QY 951 GCTGTTGAAGGTGTCATATCATGCTTTAGACACACTTCCAGAAATACATGAAACCGATA 1010
 Db 361 GCAGTTCAAAATGTTGCTGATGATGTTGATGATGCTTCTTCTGAAATACATGAAACCGATA 420
 QY 1011 TACAAATTTATTCATGATATACATACAGAAATGGAAGAATTTCTTCCAAAGGAGGGAAGA 1070
 Db 421 TACCAAGAGCTTCTCGATGTGCAACAGGAGCGGATGTGCTTCCAGAGACAGAAAG 480
 QY 1071 ACAGATCTATTTAACTGCGGCAAGAAATTTGTGAAAGAGTTTGTGAAACCTGTAGTGGT 1130
 Db 481 ACATACCGTATTTCTATCATACGAAGATATGTTTAAAGCGTACACTCGAATTTCTGAGTC 540
 QY 1131 GAAGCAAAATGGCAATGAGGG 1153
 Db 541 GAGCCGAATGGCAATGAGAGGG 563

RESULT 11
 BG645114
 DEFINITION
 EST506733 KV3 Medicago truncatula cDNA clone pkV3-39C12 5' end,
 mRNA sequence.
 BG645114
 ACCESSION
 BG645114.1 GI:13780226
 VERSION
 KEYWORDS
 EST.
 SOURCE
 Medicago truncatula (barrel medic)
 ORGANISM
 Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
 Medicago.
 1 (bases 1 to 762)
 VandenBosch,K., Endre,G., Hur,J., Beremand,P., Town,C.D.,

Bowman,C.L., Craven,M.B., Cho,J. and Fraser,C.M.
 ESTs from roots of Medicago truncatula 72 h after Rhizobium
 inoculation, 2001
 Unpublished (2001)
 Contact: VandenBosch K
 Department of Plant Biology
 University of Minnesota
 220 Biosci Center, 1445 Gortner Ave, St. Paul, MN 55108, USA
 Tel: 612 624 2755
 Fax: 612 625 1738
 Email: kvandenb@cbs.umn.edu
 M39521e TIGR sequence name: MTEBX18TK More information is
 available at: www.medicago.org
 Seq primer: SKmod (CTA gAA CTA gTg gAT CC).
 Location/Qualifiers
 1. 762
 /organism="Medicago truncatula"
 /mol_type="mRNA"
 /cultivar="genotype A17"
 /db_xref="taxon:3880"
 /clone="pkV3-39C12"
 /tissue_type="Seedling roots"
 /dev_stage="3 days post-inoculation with Sinorhizobium
 meliloti"
 /lab_host="E. coli strain XL0LR"
 /clone_lib="KV3"
 /note="Vector: pBluescript SK -; Site 1: EcoRI; Site 2:
 XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
 was directionally ligated into the Unizap XR vector from
 Stratagene and packaged using Gigapack III Gold packaging
 extracts. Plasmids containing cDNA inserts were excised
 from the recombinant lambda-Zap phage using Ex-assist
 helper phage and propagated in XL0LR cells."

ORIGIN

Query Match 14.5%; Score 239.2; DB 12; Length 762;
 Best Local Similarity 58.4%; Pred. No. 1.9e-49;
 Matches 418; Conservative 0; Mismatches 298; Indels 0; Gaps 0;

QY 544 CAAACCCCGCTCTTTTACCGAAATACAAACGGCACTAAAGCAACCCCTTTCGAAAGGT 603
 Db 47 CAAGCATCTCTCATGATATCAAGCAACATGCTTAAAGGAGACCTTCCATAAACA 106
 QY 604 TGCCAAAGATAGAGCGGCGAGTACATTCCTTTCTATCAACAACAAGATTTCTATAACA 663
 Db 107 TCCTTAGGTTGGAGCGACGAAGCTACATTTCTAGATACGAGGAAGATCTTACATAATG 166
 QY 664 AGACTTTTACTTAACTTGCTAAGTTAGAGTTCAATTTGCTTCAGTCATTTGCAAGGAAG 723
 Db 167 AAAATTTTACTCATTTTGGCAAACTAGATTTCAATATGCTCTCAAAAGTTTACATCAAAA 226
 QY 724 ACCTCAGCCATGTGTGCAAAATGGTGGAAAGCTTTTCGATATCAAGAAACGACCTTGT 783
 Db 227 AATTTGGCAACTTTTCCAAATGGTGGAAAGAGTTGGATGTTTCGTAGTAACTACCATATG 286
 QY 784 TAAGAGATAGAAATTTGTAATGCTACTTTTGGGGACTAGGTTTCAGGCTATGAGCCACAGT 843
 Db 287 CAAGAGATAGGATAGCGGAAGTTGCTTTTGGSCATTTGGGGTATATATATGAGCCCAAT 346
 QY 844 ATTCCCGGCTAGAGTTTCTTTCACAAAGCTGTTGCTGTTTAACTCTTATAGATGACA 903
 Db 347 ATTCCACGGCAAGAAAGTTATGACGAATATTTCGTCAATAATACTGTGATCGATGATA 406
 QY 904 CTTATGATCGGTATGCTTATGATGAGAACTTAACTCTTTTACTGAAGCTGTTGAAAGGT 963
 Db 407 CATATGATGATTCGGAAGAAATGATGAATTAGAATCTTTTACTAAGGCACTCGAAAGGT 466
 QY 964 GGTCAATTACATGCTTAGACACACTTCCGAATATACATGAACCGGATATACAAATTTATCA 1023
 Db 467 GGGATATTAGTTGCTTGGTAAATCTTCCGGATATACATGAAGTTTCTTACGTTGATAATCT 526
 QY 1024 TGGATACATACACAGAAATGGAAGAAATTTCTTGGAAAGGAGGAAGAACAGATCTATTTA 1083

```

Db      527 TAGATCTTTACAAAGAAATAGAGCAGGAGATGAAAAAGGAAGAGAAATATGCACITTA 586
QY      1084 ACTCGGCAAGAATTTGTGAAGAGTTTGTAGAAACCTGATGTTGAAGCAAAATGGG 1143
Db      587 ACTACTACGTGAAGAATTCATAAATACGTTCAAGCATATATAGTCTGAGGCCAGATGCT 646
QY      1144 CAAATGAGGGACACATACCAACCACTGAAGAGCATGATCCAGTTGTAATCAATTACTTGGGG 1203
Db      647 TAATGATAATATCAACCAACATAGAGAGTACATTTCGCATATCAACAGNATCATGTG 706
QY      1204 GTGCTAACCTGCTTACAAACAATTGTTATCTTGGCATGAGTGATATATTCACAAA 1259
Db      707 GTTACGCACCTGTTACAAACAATTGTTACATTTTCATGGGGATACAGCTACAGAA 762

RESULT 12
CB968730 753 bp mRNA linear EST 30-APR-2003
LOCUS CAB10002 Iva Fa_F06 Cabernet Sauvignon Flower Pre-bloom - CAB1
DEFINITION Vitis vinifera cDNA clone CAB10002_Iva_Fa_F06 5', mRNA sequence.
ACCESSION CB968730
VERSION CB968730.1 GI:30252091
KEYWORDS EST.
SOURCE Vitis vinifera
ORGANISM Vitis vinifera
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; Vitaceae; Vitis.
REFERENCE 1 (bases 1 to 753)
AUTHORS Goes da Silva,F., Randolino,A., Lim,H., Baek,J., Jones,K. and
Cook,D.
TITLE Expressed sequence tags from Vitis vinifera 'Cabernet sauvignon'
berries at various developmental stages
JOURNAL Unpublished (2003)
COMMENT Contact: Douglas Cook, PhD
CAES Genome Facility
UC Davis, Plant Pathology
One Shields Ave, Davis, CA 95616, USA
Tel: 530 754 6561
Fax: 530 754 6617
Email: drcoc@ucdavis.edu
Seq primer: ACGGTACCGGACATATGCC.
FEATURES
source
1..753
/organism="Vitis vinifera"
/mol_type="mRNA"
/cultivar="Cabernet Sauvignon"
/db_xref="taxon:29760"
/clone="CAB10002 Iva Fa_F06"
/sex="Hermaphrodite"
/dev_stage="Pre-bloom"
/lab_host="DH5alpha"
/clone_lib="Cabernet Sauvignon Flower Pre-bloom - CAB1"
/note="Organ: Flower - Pre-bloom; Vector: pDNR; Site_1:
Sfil; Site_2: Sfil; CAB1 is a cDNA library of Vitis
vinifera cv. 'Cabernet Sauvignon' clone 8 flowers. Samples
were collected approximately eleven days before onset of
bloom (clusters at this stage were fully developed and
flowers with calypters or caps still attached. Sampled
vines were located at the University of California, Davis,
Experimental Vineyard. cDNAs were made by oligo-dT priming
and directionally cloned. 5' and 3' adaptors were used in
cloning as follows:
5'-AAGCAGTGGTATCAACGAGAGTGGCCATTACGGCGGG-3' and
5'-ATTCTAGAGCGGAGCGGCGGACATG-dT(30)NN-3'. Library was
constructed using the Clontech Creator SMART kit and
size-selected to contain the 0.5-3 kb size fraction."

ORIGIN
Query Match 14.3%; Score 235.6; DB 14; Length 753;
Best Local Similarity 57.8%; Pred. No. 1.5e-48;
Matches 435; Conservative 0; Mismatches 316; Indels 1; Gaps 1;

```

```

QY      889 CTCCTATAGATGACACTTATGATCGGTATGTTATGAAGAACTTAAAGATCTTTACTG 948
Db      2 CTAATTTATGATGACATCTATGACGCGTATGTTA CACTTGAAGAACTCGAGCTCTTCACCG 61
QY      949 AAGCTGTTGAAGAGTGGTCAATTAATCATGCTTAGACACACTTCCAGAATACATGAACCGA 1008
Db      62 AAGCAGTTGAGAGTGGGATGTCACTGCCATAGATCAACTGCCAGAGTACATGAGAGTGT 121
QY      1009 TATACAAATTTATCATGGATACATACACAGAAATGGAAGATTTCTTTCGAAGGAGGAA 1068
Db      122 GTTATCAAGCACTCTTATATGATATACAGTGAATTTGAGGAAGAAATGCCCAAGGAGGAA 181
QY      1069 GAACAGATCTATTAACTCGCGCAAGAAATTTGTGAAAGAGTTTGTAGAAACCTGTATGG 1128
Db      182 GATCGTATCGGCTCTACTATGCAAAAGAGCAATGAAATCAAGTAAGACACACTATG 241
QY      1129 TTGAAGCAAAATGGGCAAAATGAGGGACACATACCAACCATGAAGACATGATCCAGTTG 1188
Db      242 AGGAAGCCAAATGGCTTCAAGTGCAACAAATTCACACCATGGAGGAGTACATGGCTGTGG 301
QY      1189 TAATCATTTACTGGCGGTCTAACTCTTACACAACTTCTTATCTTGGCATGAGTGATA 1248
Db      302 CGTTAGTTACCTCTGCATCTCAATGCTGGCGACCATCGTTTGTGGAAATGGAGATG 361
QY      1249 TATTCAAAAAGAGTCTGCGAATGGGCTGCTCTGCACTCTCTCTTTTAGATACTCAG 1308
Db      362 CTGTGACAAAAGAAACCTTCGATTGGATTTTCAGTGAGCCTAAGATTGTAAGGCTTCAG 421
QY      1309 GTATCTTGGTGACGCGCTAAATGATCTCATGCCCAAGCGCGGACGAGCAAGAAAGAAC 1369
Db      422 CCATAGTTTTCGAGGCTCATGGATGATATGTTTCCCAAAAGTTTGACGAGAAGAGAGAC 481
QY      1369 ATAGTTTCATCGAGCCTTGAAGTTTATATGAAGAAATATATGTCATGAGGAGTATGCC 1428
Db      482 ATGTTGCTTCAGCTGTGTGATGTTTATATGAAGCAACATGTCGTCGAGCAGGAAAGCG 541
QY      1429 AAACCTTGTATTTACAAAGAGTAGAAGATGTGTGGAAGATATAAAACCGAGAGTACCTCA 1488
Db      542 ACAACGAATTTAACAACAAGTAAGAGATGATGTAAGAGATATCAATGAAGAGTGCCTCA 601
QY      1489 CAACTAAAAACATTCGAAGCGCTTATGATGCTGTGATCTATTTGTGCCAGTTCTTG 1548
Db      602 TCCCAACAGCTGTCCCAATGCCATCTCTCATGGAGTTCTAAATCTCGCAGCTGTTATAG 661
QY      1549 AAGTTCAATA-TGCAGGAAGGATACTTCACACGATGGGAGACGAATACAAACATCTC 1607
Db      662 NTGTTATATACAGAATGAAGATGGGCTATCTATCTGACTGAT 1639
QY      1608 ATAAAGTCTCTACTCGTTTATCTCTATGAGTAT 1639
Db      722 GTAACCTCCATGCTAATGATCTCTGTGCCAAT 753

RESULT 13
AW685590
LOCUS NF029D03NR1F1000 Nodulated root Medicago truncatula cDNA clone
DEFINITION NF029D03NR 5', mRNA sequence.
ACCESSION AW685590
VERSION AW685590.1 GI:7560326
KEYWORDS EST.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eucosids 1; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE 1 (bases 1 to 663)
AUTHORS Watson,B.S., Shin,H.S., Lopez-Meyer,M., Scott,A.D., Harris,A.R.,
Gonzales,R.A., Bell,C.J., Flores,H.R., Inman,J.T., Weller,J.W.,
May,G.D. and Paiva,N.L.
TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation
Medicago truncatula nodulated root library

```

Unpublished (2000)
Contact: Paiva NL
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7317
Fax: 580 221 7380
Email: nlpaiva@noble.org
Insert Length: 663 Std Error: 0.00
Plate: 029 Row: D Column: 03
Seq primer: TCACAGGAACAGCTATGAC.
Location/Qualifiers
1. .663
/organism="Medicago truncatula"
/mol_type="mRNA"
/db_xref="taxon:3880"
/clone="NF029D03NR"
/tissue_type="root"
/dev_stage="Pooled developmental"
/clone_lib="Nodulated root"
/notes="Vector: Lambda Zap; Four-week old Rhizobium
meliloti-inoculated Medicago truncatula roots, containing
a mixture of young and old roots and nodules."
ORIGIN
Query Match 14.1%; Score 232.6; DB 10; Length 663;
Best Local Similarity 50.0%; Pred. No. 8.5e-48;
Matches 388; Conservative 0; Mismatches 259; Indels 0; Gaps 0;
QY 629 CATTCCCTTCTATCAACAAGATCTCTAACAAGACTTTACTTAACATGCTAAGTT 688
Db 1 CATTCTCTATGAGAAAGATCTCTCTGTGATGAAGTTCTACTCACTTTCGCAAAATT 60
QY 689 AGAGTTCAATTTGCTTCAGTCATTCGACAGGAGAGCTGACCATGTGTGCCAAATGTTG 748
Db 61 GGATTTCAATTTTACTCCATCCCTACATCAAAAAGAGTTTGGAAACATTTCCAAATGTTG 120
QY 749 GAAAGCTTCGATATCAAGAAGACGACCTGTTTGAAGATAGATGAATGTTGAATGCTA 808
Db 121 GAGGAGTTGAGCTTCTCTACCAACTTATGCAAGATAGATAGATCGCGAATGTAG 180
QY 809 CTTTGGGGACTAGTTTCAGGCTATGAGCCACAGTATTCGGGGCTAGAGTTTCTTCCAC 868
Db 181 CTTTGGGTTTGACTGCATTTTGTGAACCCCAATATTTCTCAAGCAAGAAATGATGAT 240
QY 869 ABAAGCTGTGCTGTATATCTCTATAGATGACATTTATGATCGGTATGCTATTAGA 928
Db 241 TAAAGTAATCACCTACTTTCAATCATGTATGATACATATGATCATATGGAACAATTGA 300
QY 929 AGAAGTTAAGATCTTTACTGAAGCTGTTGAAGGTGTCATTTACATGCTTAGACACT 988
Db 301 TGAAGTGAAGCTTTTACCAGGAGTTGAAGGTGGATATTTAGCAGCTTGGACGAACT 360
QY 989 TCAGAAATACATGAACCGATATACAAATTTATTCATGATACATACACAGAAATGGAAGA 1048
Db 361 TCAGACTACATGAACCTATCATATGATCATCTTTTAACTTTTACGAGAAATAGAGAA 420
QY 1049 ATTCTTGAAGAGGAGGAGACAGATCTATTATTAATCTGGCGGAAAGAAATTTGGAAGA 1108
Db 421 AGAATATGAGAAAGAGGAGGATATATACCTTGAATCTACTACAAATTTGAATTCAAAA 480
QY 1109 GTTTGTTAGAAACCTGATGTTGAGCAAAATGCGCAATGAGGACACATACCAACAC 1168
Db 481 GTCAGTGAAGCAATTTATGACAGGCGGAGATGCTGAAATGAAGAAACCATATACCAACAC 540
QY 1169 TGAAGAGCATGATCCAGTTGTAATCAATTACTTGGCGGTGCTAACTGCTTACCAACACTTG 1228
Db 541 AGAAGAGTACATGGCATATCAAAAAATCAGGTGCTTACCTTTGCTGATTTAACTC 600
QY 1229 TTATCTTGGCATGAGTATATATTCACAAAGAGTCTGTGCAATGGG 1275
Db 601 TTACATTTGGTGGTGACATAGCCACAAAGGAGATCTTCAATTTGGG 647

RESULT 14
CD485858
LOCUS
DEFINITION
CD485858 Cotton Root and Hypocotyl tissues infected with Fusarium
oxysporum f.sp. vasinfectum (Fov) Lambda ZIPLOX Library (CFUS)
Gossypium hirsutum cDNA clone CFUS3.4F02 5' similar to
(+)-delta-cadinene synthase, mRNA sequence.
CD485858
CD485858.1 GI:31406823
EST.
Gossypium hirsutum (upland cotton)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
1 (bases 1 to 777)
Dowd,C., Wilson,I. and McFadden,H.
Different Gene Expression Responses in Cotton Root and Hypocotyl
tissues during infection with Fusarium Wilt Disease
Unpublished (2003)
Contact: Caitriona Dowd, Helen McFadden
Commonwealth Scientific and Industrial Research Organisation
Division of Plant Industry
Black Mountain Laboratories, Cnr Clunies Ross Street & Barry Drive,
Black Mountain, Canberra, ACT. 2601, Australia
Tel: 61 2 6246 4914, 6246 5377
Fax: 61 2 6246 5000
Email: Caitriona.Dowd@csiro.au, Helen.McFadden@csiro.au
Vector clipped sequences Bases 1-17 (GTGACCCACGCTCG): Sali
adapter
Seq primer: M13 reverse primer
High quality sequence stop: 777.
Location/Qualifiers
1. .777
/organism="Gossypium hirsutum"
/mol_type="mRNA"
/cultivar="DeltaEMERALD"
/db_xref="taxon:3635"
/clone="CFUS3.4F02"
/tissue_type="Fov infected root and hypocotyls tissues"
/dev_stage="seedlings, pooled from 5 hours to 7 days post
infection with Fov."
/lab_host="Y1090 (ZL)"
/clone_lib="Cotton Root and Hypocotyl tissues infected
with Fusarium oxysporum f.sp. vasinfectum (Fov) Lambda
ZIPLOX Library (CFUS)"
/note="Vector: Lambda ZIPLOX; Site 1: Sali; Site 2: NotI;
mRNA was prepared from pooled root and hypocotyl tissues
of the cotton cultivar DeltaEMERALD at different stages of
Fov infection from 5 hours to 7 days post infection. cDNA
was synthesised from a NotI-oligo primer/adaptor using
the manufacturers protocols (Life Technologies) and then
ligated to a Sali adapter to facilitate directional
cloning. The cDNA was cloned into the Sali and NotI sites
of the Lambda ZIPLOX phage vector (Life Technologies).
Constructed by Caitriona Dowd and Helen McFadden."
ORIGIN
Query Match 14.0%; Score 230.6; DB 14; Length 777;
Best Local Similarity 58.3%; Pred. No. 2.8e-47;
Matches 404; Conservative 0; Mismatches 289; Indels 0; Gaps 0;
QY 705 CAGTCATTGCACAGGAAGCTCAGCCATGTGTGCAATGTGTGGAAAGCTTTCGATATC 764
Db 29 CAATTTTGCATAGAAAGAGCTAAGCGAGATTTCAGGCTGTGGAAAGATTAGACTTT 88
QY 765 AAGAGACGACCTGTTTANGAGATAGAAATGTTGATGCTACTTTGGGACTAGGT 824
Db 89 CAAGAAAGTTGCCATATGCAAGAGATAGAGTTGAAGGCTATTTTGGATCTCAGGA 148
QY 825 TCAGGCTATGACCCACAGTATTCGCGGCTAGAGTTTCTTCACAAAGCTGTTGTTGT 884

FEATURES
source

Db 149 GTGACTTTGAGCCCAATATTCACCTGGTAGAAGATGTTGACAAAGTAGCAATG 208
QY 885 ATAACTTTATAGACACTTATGATGGGTATGGTACTTATGAAACCTTAAGATCTTT 944
Db 209 GCATCTATTGATGATACATATGACTCATATGCAACATATGAAAGAGCTCATTCATAT 268
QY 945 ACTGAGCTGTTGAAAGGTGCTCAATTACATGCTTACACACACTTCCAGAAATACATGAAA 1004
Db 269 ACAAATGCAATGAGAGGTGGGATATCAAAATGATAGTGAATCTTCTGTAATACATGAAG 328
QY 1005 CGATATACAAATATTTCATGATACATACACAGAAATGGAAGAAATTTCTTGCAAAAGGAG 1064
Db 329 CCGAGCTACAAGGCACTATTAGATGTTTACGAGAAATGGAACAACTGGTGCCTGAGCAT 388
QY 1065 GGAAGAACAGATCTATTAACTGCGGCAAGAAATTTGTGAAAGAGTTTGTGAAACCTG 1124
Db 389 GGAAGACATATCTGTCGAATATGCAAAATATGCAAAATATGCGATGATACGACTTGTCTCAATCTTAC 448
QY 1125 ATGGTTGAGCAAAATGCGCAATATGAGGACACATACCAACCACTGAAGAGCATGATCCA 1184
Db 449 CTGTGGAGCCAGATGGACTCTTCAAACTACAACCATCATTCGAGGAGTTTAGGCT 508
QY 1185 GTTGAATCATTAAGTGGCGTGTAACTGCTTACAACTGTTGTTATCTTGGCATGAGT 1244
Db 509 AATGCAATGCCAATCTGTTGATGCAATGCTTGTATTACATCTTTGCGTGGCATGGGA 568
QY 1245 GATATATTCACAAAGAGTCTGCGAATGGGCTGTCTCTGACCTCTCTTTTATGATAC 1304
Db 569 GATATGTAACCAAGAGACCTTTAAATGGCGACCAATGACCTTAAGATCATTCAGCT 628
QY 1305 TCAGGTATACTTGTGTCGCGCTTAATGATCTCATGACCCACAGGCGGAGCAAGAAAGA 1364
Db 629 TCACAATTTATGTTAGTTTATGATGATGTTGCGGAACACAAAGTTTAAGCATAGGAGA 688
QY 1365 AAACATAGTTTCATCGAGCCCTTGAAGTTATPATG 1397
Db 689 GAAGACGATTGCTCAGCAATAGATGTTACATG 721

RESULT 15
BG588994
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source

1397 bp mRNA linear EST 12-APR-2001
EST490803 MHRP- Medicago truncatula cDNA clone pMHRP-59114, mRNa
sequence.
BG588994
BG588994.1 GI:13607134
EST.
Medicago truncatula (barrel medic)
Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
1 (bases 1 to 807)
Harrison,M.J.; Liu,J., Town,C.D., Van Aken,S., Utterback,T., Cho,J.
and Fraser,C.M.
ESTs from phosphate-starved roots of Medicago truncatula, 2001
Unpublished (2001)
Contact: Harrison M.J.
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73401
Tel: 580-221-5810
Fax: 580-221-7380
Email: mjharrison@noble.org
The Samuel Roberts Noble Foundation: N387745e TIGR sequence name:
MTHB55NK More information is available at: http://www.medicago.org
Seq primer: SKmcd (CTA GAA gta gat CC).
Location/Qualifiers
1. 807
/organism="Medicago truncatula"
/mol_type="mRNA"
/cultivar="A17"

/db_xref="taxon:3880"
/clones="pMHRP-59114"
/tissue_type="roots"
/dev_stage="phosphate-starved"
/lab_host="XL0LR"
/clone_lib="MHRP-"
/note="Vector: pBlueScript SK-; Site.1: EcoRI; Site.2:
XhoI; At the trifoliolate stage, M. truncatula plants were
transplanted to phosphate-free sand and grown for a
further 30 days. During this period, they were fertilized
twice weekly with 1/2 Hoaglands solutions containing 20mM
potassium phosphate. cDNA was prepared from polyA+
enriched RNA. The cDNA was directionally ligated into the
Unizap XR vector from Stratagene and packaged using
GibcoPack III Gold packaging extracts. Plasmids containing
cDNA inserts were excised from the recombinant lambda-zap
phage using Ex-assist helper phage and propagated in
XL0LR cells."

ORIGIN

Query Match 13.7%; Score 226.4; DB 12; Length 807;
Best Local Similarity 56.8%; Pred. NO. 3.3e-46;
Matches 436; Conservative 0; Mismatches 331; Indels 1; Gaps 1;
QY 883 TTATAACTTTATAGATGACACTTATGATGCTATGGTACTTTATGAAGAACTTAAGATCT 942
Db 3 TAATPACTGTGATCGATGATACATATGATGATTCGGAAGATTGATGAATTAGAACTTT 62
QY 943 TTACTGAAGCTGTTGAAAGGTGGTCAATTACATGCTTTAGACACACTTCAGAAATACATGA 1002
Db 63 TTACTAAGCACTCGAAAGGTGGGATATTAGTTGCTTGGATAATCTTCCGGATTACATGA 122
QY 1003 AACCGATATACAAATATTATTCATGATACATACAGAAATGGAAGAAATTTCTTGCAAGG 1062
Db 123 AGTTTCTCTACGTGATATCTTAGATCTTTACAAAGAAATAGAGCAGGAGATGAAAAAAG 182
QY 1063 AGGGAAGAACACAGATCTATTAACTGCGCAAGAAATTTGTGAAAGAGTTGTTAGAAACC 1122
Db 183 AAGGAGAGAAATATGCATTTACTACTAGTGAAGAAATTCATAATATAGTTCAAGCAT 242
QY 1123 TGATGGTTGAAGCAAAATGCGCAATATGAGGGACACATACCAACCACTGAAGAGCATGATC 1182
Db 243 ATATGACTGAGGCCAGATGGTTAAATGATAAATATCAACCAACATTTAGAAGAGTACATTC 302
QY 1183 CAGTTGTAATCATTTACTGCGGTGCTAACCTGCTTACACAACTGTTTATCTGGCATGA 1242
Db 303 GCATATCAACAGAAATCATGTGGTACGCACTGGTTACACAACTGTTTACATTTTCATGG 362
QY 1243 GTGATATATTCACAAAGAGTCTGTGGAATGGGTGCTCTCTGCACCTCTCTTTTATAGAT 1302
Db 363 GGGATACAGCTACAGAGACATCTTCAATGGGTATCAATGGGCCAAATAATCATCAATG 422
QY 1303 ACTCAGGTATATCTTGGTCAAGCCCTAAATGATCTCATGACCCCAAGGCGGAGCAAGAAA 1362
Db 423 CTGCTATTATCTTTGCGAGGCTAATGGATGATATTGTCATCCACAGAGTTTGAACAAAAA 482
QY 1363 GAAACATATGTTTCAGAGCCCTTGAAGTTATATGAGAAATATATGTCATGAGGAGT 1422
Db 483 GAGATCATGTTTCTCTCATTTTGGAAATGTTATATGAAGCAACATAATATCTCTAGAGAAG 542
QY 1423 ATGCCCAACCTTGTATTTACAGGAAGTAGAAGATGTGTGGAAGATATAAACCAGAGT 1482
Db 543 GTGCCATTCAGAGAGTTCGAAAAAGAAATTTGTGATGCTTGGAAAGATATGAATAAGAAAT 602
QY 1483 ACCTCACAACTAAAAACATTCGAGCCGCTTATGATGGCTGTGATCTATTGTTCGCAAGT 1542
Db 603 GTCTTATCCGACTGAAATTCACCTCTTTTGGACGTGCACTTTAAACCTTATACGTT 662
QY 1543 TTCTTGAAGTTCAATATGAGGAAAGGATACTTACAGTATGGGAGGAGATACAAAC 1602
Db 663 TCATGGATGTGTTTATAAAGATAAAGATAACTTACACATCCCTGAGGAGAAATGAAGA 722
QY 1603 ATCTCATAAAGTCTCT-ACTCGTTTATCTCTATGATGATGAGGATCC 1649

Db 723 CATTCATCAAAATCCTTTAACTTGTGATCCAGTGCCCAATTGAGCATAC 770

Search completed: June 7, 2004, 12:38:38
Job time : 4510 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2004, 16:37:26 ; Search time 45 Seconds
(without alignments)
3849.322 Million cell updates/sec

Title: US-09-763-822A-14
Perfect score: 2906
Sequence: 1 TMAITEKPIRPIANFPPI.....GDEYKHLIKSLVPMISGS 549

Scoring table: BLOSUM62
Gapop 10.0 ; Gapext 0.5

Searched: 1017041 seqs, 315518202 residues
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL_25.*
- 1: sp_archaea.*
 - 2: sp_bacteria.*
 - 3: sp_fungi.*
 - 4: sp_human.*
 - 5: sp_invertebrate.*
 - 6: sp_mammal.*
 - 7: sp_mhc.*
 - 8: sp_organelle.*
 - 9: sp_phage.*
 - 10: sp_plant.*
 - 11: sp_rodent.*
 - 12: sp_virus.*
 - 13: sp_vertebrate.*
 - 14: sp_unclassified.*
 - 15: sp_rvirus.*
 - 16: sp_bacteriap.*
 - 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match	Length	ID	Description
1	2888	99.4	546	10 Q9FY41	Q9fy41 artemia a
2	2885	99.3	546	10 Q9FVMS	Q9fvm5 artemia a
3	2878	99.0	546	10 Q9LW98	Q9lww8 artemia a
4	2864	98.6	546	10 Q9AR04	Q9ar04 artemia a
5	1730	59.5	548	10 Q9SA63	Q9sa63 artemia a
6	1699.5	58.5	549	10 Q9FYX6	Q9fxy6 artemia a
7	1577.5	54.3	549	10 Q9AR67	Q9ar67 solidago ca
8	1463.5	50.4	547	10 Q9ST45	Q9st45 artemia a
9	1462.5	50.3	547	10 Q9LLR9	Q9llr9 artemia a
10	1314.5	45.2	573	10 Q9AXP5	Q9axp5 artemia a
11	1303.5	44.9	577	10 Q9FXV7	Q9fxv7 artemia a
12	1183	40.7	544	10 Q9FQ27	Q9fq27 lycopersico
13	1156	39.8	546	10 Q9FQ26	Q9fq26 lycopersico
14	1143	39.3	555	10 Q9SAN0	Q9san0 gossypium a
15	1139.5	39.2	555	10 Q9FYU6	Q9fyu6 citrus juno
16	1135	39.1	548	10 Q64961	Q64961 lycopersico

17	1130	38.9	551	10 Q9LKN1	Q9lkn1 gossypium h
18	1130	38.9	552	10 Q9FQ28	Q9fq28 lycopersico
19	1123	38.6	548	10 Q64962	Q64962 lycopersico
20	1092	37.6	548	10 Q8RVR2	Q8rvr2 citrus para
21	1063.5	36.6	551	10 Q9ZTQ6	Q9ztq6 solanum tub
22	1060	36.5	550	10 Q9SBJ0	Q9sbj0 solanum tub
23	1059	36.4	556	10 Q9ZTQ7	Q9ztq7 solanum tub
24	1056	36.3	555	10 Q84LK9	Q84lk9 gossypium a
25	1055	36.3	549	10 Q9XJ25	Q9xj25 solanum tub
26	1041	35.8	556	10 Q9XJ32	Q9xj32 solanum tub
27	1040.5	35.8	557	10 Q9ZTQ8	Q9ztq8 solanum tub
28	1039	35.8	549	10 Q9XIZ0	Q9xiz0 solanum tub
29	1038	35.7	598	10 Q94EN2	Q94en2 nicotiana t
30	1037.5	35.7	556	10 Q9FVL3	Q9fvl3 lycopersico
31	1036.5	35.7	559	10 Q81923	Q81923 capsicum an
32	1032.5	35.5	559	10 Q65323	Q65323 capsicum an
33	1028.5	35.4	553	10 Q9ATN6	Q9atn6 capsicum an
34	1020	35.1	560	10 Q9SDN9	Q9sdn9 capsicum an
35	1017	35.0	597	10 Q94IH1	Q94ih1 nicotiana t
36	1012.5	34.8	548	10 Q84LF0	Q84lf0 nicotiana a
37	1009.5	34.7	548	10 Q84LF2	Q84lf2 nicotiana a
38	1005.5	34.6	561	10 Q81634	Q81634 elaeis olei
39	999.5	34.4	560	10 Q94US8	Q94us8 citrus juno
40	996.5	34.3	548	10 Q84LGO	Q84lgo nicotiana a
41	992	34.1	508	10 Q9SW77	Q9sw77 gossypium a
42	988.5	34.0	558	10 Q8LSC2	Q8lsc2 cichorium i
43	984.5	33.9	559	10 Q8S3A6	Q8s3a6 lactuca eat
44	983.5	33.8	548	10 Q84LF1	Q84lf1 nicotiana a
45	978	33.7	520	10 Q39978	Q39978 hyoscyamus

ALIGNMENTS

RESULT 1

Q9FY41 PRELIMINARY; PRT; 546 AA.
ID Q9FY41
AC Q9FY41;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Amorpha-4,11-diene synthase.
OS Artemisia annua (Sweet wormwood).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC campanulids; Asterales; Asteraceae; Asteraceae; Anthemideae;
OC Artemisia.
OX NCBI_TaxID=35608;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21183247; PubMed=11289612;
RA Wallaart T.E., Bouwmeester H.J., Hille J., Poppinga L., Majers N.C.;
RT "Amorpha-4,11-diene synthase: cloning and functional expression of a
RT key enzyme in the biosynthetic pathway of the novel antimalarial drug
RT artemisinin";
RL Planta 212:460-465(2001).
DR EMBL; AY006482; AAF98444.1; --
DR HSP; Q40577; SEAU.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR005630; Terpene synth C.
DR InterPro; IPR008949; Terpenoid synth.
DR InterPro; IPR008930; Terp_cycToroid.
DR InterPro; IPR001906; Terp_synth-like.
DR Pfam; PF01397; Terpene synth; 1.
DR Pfam; PF03936; Terpene synth C; 1.
SQ SEQUENCE 546 AA; 63532 MW; FCD786A7ED81FB21 CRC64;

Query Match 99.4%; Score 2888; DB 10; Length 546;
Best Local Similarity 99.8%; Pred.No. 3.1e-206;
Matches 545; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 MALTEKPIRPIANFPPI...GDEYKHLIKSLVPMISGS 549 QY 2 MALTEKPIRPIANFPPI...GDEYKHLIKSLVPMISGS 549

Db 1 MSLTEKPIRPIANFPFSGDQFLIYQKQVQGVQVNDLKKEVRQLLKEALDIPMKH 60
QY 62 ANLLKLIIDEIQRGLGPIYHFEREIDHALQCIYETYGDNWNGDRSSLWFLMRKQGYVTCD 121
Db 61 ANLLKLIIDEIQRGLGPIYHFEREIDHALQCIYETYGDNWNGDRSSLWFLMRKQGYVTCD 120
QY 122 VFNNYKDKNGAFKQSLANDVEGLLELYEATSMRVPGEIILEDALGFTSRSLRSLMTKDAFS 181
Db 121 VFNNYKDKNGAFKQSLANDVEGLLELYEATSMRVPGEIILEDALGFTSRSLRSLMTKDAFS 180
QY 182 TNPALFTIQRALKQPLWKLRIEAAQYIPFYQQQDSHNKTLKLAKEFNLLQSLHKE 241
Db 181 TNPALFTIQRALKQPLWKLRIEAAQYIPFYQQQDSHNKTLKLAKEFNLLQSLHKE 240
QY 242 ELSHVCKWKAFAFIKKNAPCLDRIVECYFWGLSGGYEPOYSRVRVFTKAVAVITLIDD 301
Db 241 ELSHVCKWKAFAFIKKNAPCLDRIVECYFWGLSGGYEPOYSRVRVFTKAVAVITLIDD 300
QY 302 TYDAYGTGYEELKIPTFAVERWSITCLDTLPEYMKPIYKLFMDTYTMEBEFLAKEGRDTLF 361
Db 301 TYDAYGTGYEELKIPTFAVERWSITCLDTLPEYMKPIYKLFMDTYTMEBEFLAKEGRDTLF 360
QY 362 NCGKEFVKFVFNLMVAKWANEHGPTEEBHPVVIITGGANLLTTTCYLGMSDIFTKE 421
Db 361 NCGKEFVKFVFNLMVAKWANEHGPTEEBHPVVIITGGANLLTTTCYLGMSDIFTKE 420
QY 422 SVWAVSAPPLFRYSGLGRRLNDLMTKABQERKSSSLESYMKYNNWEEYAQTLLY 481
Db 421 SVWAVSAPPLFRYSGLGRRLNDLMTKABQERKSSSLESYMKYNNWEEYAQTLLY 480
QY 482 KEVEDVWKDINREYLTNNIPRPLMAVIYLCQFLEVQYAGKDNFTRMGDEYKHLIKSL 541
Db 481 KEVEDVWKDINREYLTNNIPRPLMAVIYLCQFLEVQYAGKDNFTRMGDEYKHLIKSL 540
QY 542 VYPMSI 547
Db 541 VYPMSI 546

RESULT 2
Q9FWMS PRELIMINARY; PRT; 546 AA.
ID Q9FWMS PRELIMINARY; PRT; 546 AA.
AC Q9FWMS;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Amorpha-4,11-diene synthase.
OS Artemisia annua (Sweet wormwood).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC campanulids; Asterales; Asteraceae; Asteroideae; Anthemideae;
OC Artemisia.
OX NCBI_TaxID=35608;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RX MEDLINE=20485153; PubMed=11032404;
RA Merck P.E., Bengtsson M., Boumeester H.J., Posthumus M.A.,
RA Brodelius P.E.;
RT "Molecular cloning, expression, and characterization of a amorpho-
4,11-diene synthase from, a key enzyme of artemisinin biosynthesis of
Artemisia annua L.";
RL Arch. Biochem. Biophys. 381:173-180(2000).
DR EMBL; AF138959; RA61439.1; --
DR HSSP; Q40577; SEAU.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR005630; Terpene synth C.
DR InterPro; IPR008949; Terpenoid synth.
DR InterPro; IPR008930; Terp_cyc_Loxid.
DR InterPro; IPR001906; Terp_synth-like.
DR Pfam; PF01397; Terpene_synth; 1.

DR Pfam; PF03936; Terpene synth_C; 1.
SQ SEQUENCE 546 AA; 63933 MW; 381386A7ED81FEF9 CRC64;
Query Match 99.3%; Score 2885; DB 10; Length 546;
Best Local Similarity 99.6%; Pred. No. 5.1e-206;
Matches 544; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 2 MALTEEKPIRPIANFPFSGDQFLIYQKQVQGVQVNDLKKEVRQLLKEALDIPMKH 61
Db 1 MSLTEKPIRPIANFPFSGDQFLIYQKQVQGVQVNDLKKEVRQLLKEALDIPMKH 60
QY 62 ANLLKLIIDEIQRGLGPIYHFEREIDHALQCIYETYGDNWNGDRSSLWFLMRKQGYVTCD 121
Db 61 ANLLKLIIDEIQRGLGPIYHFEREIDHALQCIYETYGDNWNGDRSSLWFLMRKQGYVTCD 120
QY 122 VFNNYKDKNGAFKQSLANDVEGLLELYEATSMRVPGEIILEDALGFTSRSLRSLMTKDAFS 181
Db 121 VFNNYKDKNGAFKQSLANDVEGLLELYEATSMRVPGEIILEDALGFTSRSLRSLMTKDAFS 180
QY 182 TNPALFTIQRALKQPLWKLRIEAAQYIPFYQQQDSHNKTLKLAKEFNLLQSLHKE 241
Db 181 TNPALFTIQRALKQPLWKLRIEAAQYIPFYQQQDSHNKTLKLAKEFNLLQSLHKE 240
QY 242 ELSHVCKWKAFAFIKKNAPCLDRIVECYFWGLSGGYEPOYSRVRVFTKAVAVITLIDD 301
Db 241 ELSHVCKWKAFAFIKKNAPCLDRIVECYFWGLSGGYEPOYSRVRVFTKAVAVITLIDD 300
QY 302 TYDAYGTGYEELKIPTFAVERWSITCLDTLPEYMKPIYKLFMDTYTMEBEFLAKEGRDTLF 361
Db 301 TYDAYGTGYEELKIPTFAVERWSITCLDTLPEYMKPIYKLFMDTYTMEBEFLAKEGRDTLF 360
QY 362 NCGKEFVKFVFNLMVAKWANEHGPTEEBHPVVIITGGANLLTTTCYLGMSDIFTKE 421
Db 361 NCGKEFVKFVFNLMVAKWANEHGPTEEBHPVVIITGGANLLTTTCYLGMSDIFTKE 420
QY 422 SVWAVSAPPLFRYSGLGRRLNDLMTKABQERKSSSLESYMKYNNWEEYAQTLLY 481
Db 421 SVWAVSAPPLFRYSGLGRRLNDLMTKABQERKSSSLESYMKYNNWEEYAQTLLY 480
QY 482 KEVEDVWKDINREYLTNNIPRPLMAVIYLCQFLEVQYAGKDNFTRMGDEYKHLIKSL 541
Db 481 KEVEDVWKDINREYLTNNIPRPLMAVIYLCQFLEVQYAGKDNFTRMGDEYKHLIKSL 540
QY 542 VYPMSI 547
Db 541 VYPMSI 546

RESULT 3
Q9LW98 PRELIMINARY; PRT; 546 AA.
ID Q9LW98 PRELIMINARY; PRT; 546 AA.
AC Q9LW98;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Amorpha-4,11-diene synthase.
GN KCS12.
OS Artemisia annua (Sweet wormwood).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC campanulids; Asterales; Asteraceae; Asteroideae; Anthemideae;
OC Artemisia.
OX NCBI_TaxID=35608;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21036108; PubMed=11185551;
RA Chang Y.U., Song S.H., Park S.H., Kim S.U.;
RT "Amorpha-4,11-diene synthase of Artemisia annua: cDNA isolation and
RT bacterial expression of a terpene synthase involved in artemisinin
RT biosynthesis.";
RL Arch. Biochem. Biophys. 383:178-184(2000).
DR EMBL; AJ251751; CAB94691.1; --
DR HSSP; Q40577; SEAT.

DR	GO: 0016929; F: lyase activity; IEA.
DR	GO: 0008152; P: metabolism; IEA.
DR	InterPro: IPR005630; Terpene synth C.
DR	InterPro: IPR008949; Terpenoid synth.
DR	InterPro: IPR008930; Terp_cyc toroid.
DR	InterPro: IPR001906; Terp_synth-like.
DR	Pfam: PF01397; Terpene synth; 1.
DR	Pfam: PF03936; Terpene synth; C; 1.
SQ	SEQUENCE 546 AA; 63935 MW; 981A0F9B39876464 CRC64;
Query Match 99.0%; Score 2878; DB 10; Length 546;	
Best Local Similarity 99.3%; Pred. No. 1.7e-205; Indels 0; Gaps 0;	
Matches 542; Conservative 4; Mismatches 0;	
QY	2 MALTEEKPIRPIANFPSPISWGDQFLIYKQVQGVQVEIVNDLKKEVRQLLKEALDIPMKH 61
DB	1 MSLTEKPIRPIANFPSPISWGDQFLIYKQVQGVQVEIVNDLKKEVRQLLKEALDIPMKH 60
QY	62 ANLLKLDEIQRLGIPYHFERIDHALQCIYETYGDNWNGDRSSLWFLMRKQGYVTCD 121
DB	61 ANLLKLDEIQRLGIPYHFERIDHALQCIYETYGDNWNGDRSSLWFLMRKQGYVTCD 120
QY	122 VFNNYKDKNGAFKQSLANDVGLLEVEATSMRVPGEIILEDALGFTSRSLMTKDAPS 181
DB	121 VFNNYKDKNGAFKQSLANDVGLLEVEATSMRVPGEIILEDALGFTSRSLMTKDAPS 180
QY	182 TNPALFTEIQRLKQPLWKRLPRIEAAQYIPFYQQQDSHNKTLKLAKLEFNLLQSLHKE 241
DB	181 TNPALFTEIQRLKQPLWKRLPRIEAAQYIPFYQQQDSHNKTLKLAKLEFNLLQSLHKE 240
QY	242 ELSHVCKWKWAFDIKKNAPCLDRIVECYFWGLSGYEPQYSRVRVFTKAVAVITLIDD 301
DB	241 ELSHVCKWKWAFDIKKNAPCLDRIVECYFWGLSGYEPQYSRVRVFTKAVAVITLIDD 300
QY	302 TYDAYGTVEELKIIFTEAVERWSITCLDTLPYMKPIYKLFMDTYTEMEEFLLAKEGRTDLF 361
DB	301 TYDAYGTVEELKIIFTEAVERWSITCLDTLPYMKPIYKLFMDTYTEMEEFLLAKEGRTDLF 360
QY	362 NCGKEFKVKEFVRNLMVEAKWANEHGIPTTEEHDPVVIITGGANLLTTTCYLGMSDIFTKE 421
DB	361 NCGKEFKVKEFVRNLMVEAKWANEHGIPTTEEHDPVVIITGGANLLTTTCYLGMSDIFTKE 420
QY	422 SVEWAVSAPPLFRYSGLIGRLRLNDLMTWKAEQERKSSSSLSYMKYNNVEEYATLIY 481
DB	421 SVEWAVSAPPLFRYSGLIGRLRLNDLMTWKAEQERKSSSSLSYMKYNNVEEYATLIY 480
QY	482 KEVEDVWKDINREYLTITKNIPRPLLMVAVIYLCQFLEVQYAGKDNFTRMGDYKHLIKSL 541
DB	481 KEVEDVWKDINREYLTITKNIPRPLLMVAVIYLCQFLEVQYAGKDNFTRMGDYKHLIKSL 540
QY	542 VYPMSI 547
DB	541 VYPMSI 546
RESULT 4	
Q9AR04 PRELIMINARY; PRT; 546 AA.	
AC	Q9AR04;
DT	01-JUN-2001 (Tremblrel. 17, Created)
DT	01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT	01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE	Amorpha-4,11-diene synthase.
GN	AMS1
OS	Artemisia annua (Sweet wormwood).
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC	campanulids; Asterales; Asteraceae; Asteroideae; Anthemideae;
OC	Artemisia
OX	NCBI_TaxID=35608;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=001; TISSUE=Leaf;

RA	Liu Y., Ye H.C., Li G.F.;
RT	"Cloning of sesquiterpene cyclase gene from Artemisia annua.";
RL	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN	[2]
RP	SEQUENCE FROM N.A.
RC	STRAIN=001; TISSUE=Leaf;
RA	Liu Y., Ye H.C., Li G.F.;
RT	"Cloning of sesquiterpene cyclase cDNA from Artemisia annua.";
RL	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR	EMBL: AF327527; AAK15697.1; -
DR	EMBL: AF327526; AAK15696.1; -
DR	HSSP; Q40577; SEAU.
DR	GO: 0016829; F: lyase activity; IEA.
DR	GO: 0008152; P: metabolism; IEA.
DR	InterPro: IPR005630; Terpene synth C.
DR	InterPro: IPR008949; Terpenoid synth.
DR	InterPro: IPR008930; Terp_cyc toroid.
DR	InterPro: IPR001906; Terp_synth-like.
DR	Pfam: PF01397; Terpene synth; 1.
DR	Pfam: PF03936; Terpene synth; C; 1.
SQ	SEQUENCE 546 AA; 63909 MW; 78572A680D008961 CRC64;
Query Match 98.6%; Score 2864; DB 10; Length 546;	
Best Local Similarity 98.7%; Pred. No. 1.9e-204; Indels 0; Gaps 0;	
Matches 539; Conservative 7; Mismatches 0;	
QY	2 MALTEEKPIRPIANFPSPISWGDQFLIYKQVQGVQVEIVNDLKKEVRQLLKEALDIPMKH 61
DB	1 MSLTEKPIRPIANFPSPISWGDQFLIYKQVQGVQVEIVNDLKKEVRQLLKEALDIPMKH 60
QY	62 ANLLKLDEIQRLGIPYHFERIDHALQCIYETYGDNWNGDRSSLWFLMRKQGYVTCD 121
DB	61 ANLLKLDEIQRLGIPYHFERIDHALQCIYETYGDNWNGDRSSLWFLMRKQGYVTCD 120
QY	122 VFNNYKDKNGAFKQSLANDVGLLEVEATSMRVPGEIILEDALGFTSRSLMTKDAPS 181
DB	121 VFNNYKDKNGAFKQSLANDVGLLEVEATSMRVPGEIILEDALGFTSRSLMTKDAPS 180
QY	182 TNPALFTEIQRLKQPLWKRLPRIEAAQYIPFYQQQDSHNKTLKLAKLEFNLLQSLHKE 241
DB	181 TNPALFTEIQRLKQPLWKRLPRIEAAQYIPFYQQQDSHNKTLKLAKLEFNLLQSLHKE 240
QY	242 ELSHVCKWKWAFDIKKNAPCLDRIVECYFWGLSGYEPQYSRVRVFTKAVAVITLIDD 301
DB	241 ELSHVCKWKWAFDIKKNAPCLDRIVECYFWGLSGYEPQYSRVRVFTKAVAVITLIDD 300
QY	302 TYDAYGTVEELKIIFTEAVERWSITCLDTLPYMKPIYKLFMDTYTEMEEFLLAKEGRTDLF 361
DB	301 TYDAYGTVEELKIIFTEAVERWSITCLDTLPYMKPIYKLFMDTYTEMEEFLLAKEGRTDLF 360
QY	362 NCGKEFKVKEFVRNLMVEAKWANEHGIPTTEEHDPVVIITGGANLLTTTCYLGMSDIFTKE 421
DB	361 NCGKEFKVKEFVRNLMVEAKWANEHGIPTTEEHDPVVIITGGANLLTTTCYLGMSDIFTKE 420
QY	422 SVEWAVSAPPLFRYSGLIGRLRLNDLMTWKAEQERKSSSSLSYMKYNNVEEYATLIY 481
DB	421 SVEWAVSAPPLFRYSGLIGRLRLNDLMTWKAEQERKSSSSLSYMKYNNVEEYATLIY 480
QY	482 KEVEDVWKDINREYLTITKNIPRPLLMVAVIYLCQFLEVQYAGKDNFTRMGDYKHLIKSL 541
DB	481 KEVEDVWKDINREYLTITKNIPRPLLMVAVIYLCQFLEVQYAGKDNFTRMGDYKHLIKSL 540
QY	542 VYPMSI 547
DB	541 VYPMSI 546
RESULT 5	
Q8SA63 PRELIMINARY; PRT; 548 AA.	
ID	Q8SA63
AC	Q8SA63;
DT	01-JUN-2002 (Tremblrel. 21, Created)
DT	01-JUN-2002 (Tremblrel. 21, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Beta-carboxyphenyl synthase QHS1.
GN QHS1.
OS Artemisia annua (Sweet wormwood).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC campanulids; Asterales; Asteraceae; Asteroidae; Anthemideae;
OC Artemisia.
OX NCBI_TaxID=35608;
RN [1] SEQUENCE FROM N.A.
RA Cai Y., Crock J., Jia J.-W., Lin Z.-X., Croteau R., Chen X.-Y.;
RT "Beta-Carboxyphenyl Synthase From Artemisia annua L.: cDNA Isolation,
RT Characterization, Wound Induction and Elicitation.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF472361; AAL79181.1; -;
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR005630; Terpene synthase C.
DR InterPro; IPR008949; Terpenoid synth.
DR InterPro; IPR008930; Terp_cyc_foroid.
DR InterPro; IPR001906; Terp synth-like.
DR Pfam; PF01397; Terpene synth; 1.
DR Pfam; PF03936; Terpene synth_C; 1.
SQ SEQUENCE 548 AA; 63737 MW; 40B8E69751FF6654 CRC64;
Query Match 59.5%; Score 1730; DB 10; Length 548;
Best Local Similarity 58.0%; Pred. No. 3.9e-120;
Matches 318; Conservative 101; Mismatches 127; Indels 2; Gaps 2;
QY 2 MALTEKPIRTPANPPSPWQDFLIY-QKQVEQ-GVEQIVNDLKKEVROLLKEALDIPM 59
DB 1 MSVKEEKVIRPIVHFPPSWADQFLIPDDKQAEQANVEQVNVNELREDVRKDLVSSLDVQT 60
QY 60 KHANLLKLDEIQRLGIPYHFEREIDHALQCIYETYGDNWNGDRSSLMFLMRKQGYVT 119
DB 61 EHTNLLKLDAIQRLGIAHYFEEIEQALQIHYDTYGDWKGSPSLMFLMRKQGYVYS 120
QY 120 CDVFNKYKDKNGAFKQSLANDVEGLLELYEATSRVPEGEIILBDALGFTSRSLSMTKDA 179
DB 121 CDIFPNKYKEDGSFRESLNDVEGLLELYEATYLRVQGEGLDLDALVFTTCLEKAKDL 180
QY 180 FSTNPALFTEIQRLAKQPLWKRLPRIEAAQYIPFYQQQDSHNKTLKLAKLEFNLLQSLH 239
DB 181 VHTNPTLSYIQEALQPLKRLTRLEALRYIPMYEQQAASHNESLLKLAKLGFNLLQSLH 240
QY 240 KEELSHVCKWKAFFDKKNAAPCLDRIVECYFGLGSGYEPQYSRARPVFTKAVITLI 299
DB 241 RKELSEVSRWVKGLDVFNPLPYARDRMVECYFVALGVYFEPKYSQARIFLAKVISLATVL 300
QY 300 DDTYDAYGYEELKIFTEAVERWSITCLDTPYMKPIYKLFMDTYTEMEFLAKEGRTD 359
DB 301 DDTYDAYGYEELKIFTEAVERWSITCLDTPYMKPIYKLFMDTYTEMEFLAKEGRTD 360
QY 360 LENCCKEYFVFNLMVAKWANEHGIPTTEHDPVVIITGGANLLTTTCYLGMSDIFT 419
DB 361 HLSYAKESKMEFIRSYMMEAKWANEHGVPTAEHMSVAFVSSGYSLMATTTCFVGMGDIVT 420
QY 420 KESVWAVSAPPLFRYSGLGRNLDMTHKABQERKHSSSLSSEYKMEYNNVEEYAQTL 479
DB 421 DDAFWALTKPPIIKASCAIARLMDIHQSKEEKERIHVASSVESYKQYDVTEHVLKV 480
QY 480 IYKEVEDVWKDINREYLTNNIPRLLMAVILYLCQFLEVQYAGKDNFTRMGDEYKHLIKS 539
DB 481 FNKKIEDAWKDTRESLVKDKIPMLMVRVINLAQVMDVLYKHKGDTNVGEELKDHIKS 540
QY 540 LLVYPMWI 547
DB 541 LLVHPPI 548

RESULT 6
Q9FYX6

ID Q9FYX6 PRELIMINARY; PRT; 549 AA.
AC Q9FYX6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative sesquiterpene cyclase.
GN CASC34.
OS Artemisia annua (Sweet wormwood).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC campanulids; Asterales; Asteraceae; Asteroidae; Anthemideae;
OC Artemisia.
OX NCBI_TaxID=35608;
RN [1] SEQUENCE FROM N.A.
RA Van Geldre E.G., De Pauw I., Inze D., Van Montagu M.,
RT "Cloning and molecular analysis of two new sesquiterpene cyclases from
RT Artemisia annua L.";
RL Plant Sci. 158:163-171(2000).
DR EMBL; AJ271793; CAC12732.1; -;
DR HSPF; Q40577; SEAR.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR005630; Terpene synthase C.
DR InterPro; IPR008949; Terpenoid synth.
DR InterPro; IPR008930; Terp_cyc_foroid.
DR InterPro; IPR001906; Terp synth-like.
DR Pfam; PF01397; Terpene synth; 1.
DR Pfam; PF03936; Terpene synth_C; 1.
SQ SEQUENCE 549 AA; 63808 MW; DADBE5C8244BACFE CRC64;

Query Match 58.5%; Score 1699.5; DB 10; Length 549;
Best Local Similarity 57.2%; Pred. No. 7.3e-118;
Matches 314; Conservative 101; Mismatches 131; Indels 3; Gaps 3;

QY 2 MALTEKPIRTPANPPSPWQDFLIY-QKQVEQ-GVEQIVNDLKKEVROLLKEALDIPM 59
DB 1 MSVKEEKVIRPIVHFPPSWADQFLIPDDKQAEQANVEQVNVNELREDVRKDLVSSLDVQT 60
QY 60 KHANLLKLDEIQRLGIPYHFEREIDHALQCIYETYGDNWNGDRSSLMFLMRKQGYVT 119
DB 61 EHTNLLKLDAIQRLGIAHYFEEIEQALQIHYDTYGDWKGSPSLMFLMRKQGYVYS 120
QY 120 CDVFNKYKDKNGA-FKQSLANDVEGLLELYEATSRVPEGEIILBDALGFTSRSLSMTKD 178
DB 121 CDIFPNKYKEDGSFRESLNDVEGLLELYEATYLRVQGEGLDLDALVFTTCLEKAKDL 180
QY 179 AFSTNPALFTEIQRLAKQPLWKRLPRIEAAQYIPFYQQQDSHNKTLKLAKLEFNLLQSL 238
DB 181 LVHTNPTLSYIQEALQPLKRLTRLEALRYIPMYEQQAASHNESLLKLAKLGFNLLQSL 240
QY 239 HKELSHVCKWKAFFDKKNAAPCLDRIVECYFGLGSGYEPQYSRARPVFTKAVITLI 298
DB 241 HKELSEVSRWVKGLDVFNPLPYARDRMVECYFVALGVYFEPKYSQARIFLAKVISLATV 300
QY 299 IDDTYDAYGYEELKIFTEAVERWSITCLDTPYMKPIYKLFMDTYTEMEFLAKEGRT 358
DB 301 LDDTYDAYGYEELKIFTEAVERWSITCLDTPYMKPIYKLFMDTYTEMEFLAKEGRT 360
QY 359 DLFCNGKEFVFNLMVAKWANEHGIPTTEHDPVVIITGGANLLTTTCYLGMSDIFT 418
DB 361 HLSYAKESKMEFIRSYMMEAKWANEHGVPTAEHMSVAFVSSGYSLMATTTCFVGMGDIV 420
QY 419 KESVWAVSAPPLFRYSGLGRNLDMTHKABQERKHSSSLSSEYKMEYNNVEEYAQTL 478
DB 421 DDAFWALTKPPIIKASCAIARLMDIHQSKEEKERIHVASSVESYKQYDVTEHVLKV 480
QY 479 IYKEVEDVWKDINREYLTNNIPRLLMAVILYLCQFLEVQYAGKDNFTRMGDEYKHLIK 538
DB 481 VFNKKIEDAWKDTRESLVKDKIPMLMVRVINLAQVMDVLYKHKGDTNVGEELKDHIK 540


```

Db 361 NYIKETLKEAVQSYMTAEARWAKBEYIPTIEHTKVSYSIGYKLALVAGFACMGDVIADD 420
QY 422 SVEWAVSAPPLFRYSIGILGRRLNDLMTKAEORKHSSSSLESYMKYNNVEEVAQTLLY 481
Db 421 SFWEVFNPLVACCLLCRTMDLGLSHKGEQDRKHVASTIECYMKQPDASEQAYESLN 480
QY 482 KEVEDVKDINREY-LTTKNIPRLMAVYLCQFLEVOYAGKDNFTRMGDYKHLIKSL 540
Db 481 KKVEDAWKEINREFMITCKDNIHVAMRVLPNSRSDVLYKNDHFTHVGVVEVINHIKSL 540
QY 541 LV 542
Db 541 FV 542

RESULT 9
Q9ALLR9 PRELIMINARY; PRT; 547 AA.
ID Q9ALLR9 PRELIMINARY; PRT; 547 AA.
AC Q9ALLR9;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE 8-epicetrol synthase (Epi-cetrol synthase).
GN ECSI.
OS Artemisia annua (Sweet wormwood).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC campanulids; Asterales; Asteraceae; Asteroidae; Anthemideae;
OC Artemisia.
OX NCBI_TaxID=35608;
RN [1]
RP SEQUENCE FROM N.A.
RA Hua L., Matsuda S.P.T.;
RX TISSUE=Leaf;
RM MEDLINE=99417501; PubMed=10486140;
RA Mercke P., Xue Z.T., Brodelius P.E.;
RT "Isolation, characterization and expression of an Artemisia annua cDNA
RL encoding a terpene cyclase.";
RL Arch. Biochem. Biophys. 369:213-222(1999).
DR EMBL; AF157059; AAF80333.1; -
DR EMBL; AJ001539; CAC08805.1; -
DR HSSP; Q40577; SEAT.
DR GO; GO:0016829; P:lyase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR005630; Terpene synth C.
DR InterPro; IPR008949; Terpenoid synth.
DR InterPro; IPR008930; Terp_cyc_Toroid.
DR InterPro; IPR001906; Terp_synth-like.
DR Pfam; PF01397; Terpene synth; 1.
DR Pfam; PF03936; Terpene synth C; 1.
SQ SEQUENCE 547 AA; 63564 MW; E84EE9FCEED70BE2 CRC64;

Query Match 50.3%; Score 1462.5; DB 10; Length 547;
Best Local Similarity 52.0%; Pred. No. 3e-100;
Matches 282; Conservative 97; Mismatches 162; Indels 1; Gaps 1;

QY 2 MALTEKPIRPIANFPSPISWGQFLYKQVQGVQVQVNDLKEVRLKEALDIPMKH 61
Db 1 MSLIVEDVIRPNANFPSPISWGQFLAYDQDEQGVQVQVQVNDLKEVRLKEALDIPMKH 60
QY 62 ANLLKLIDBIQRLGPIHFEREIDHALQCTIYETIYGNWNGDRSLFRLNKGQVYVTC 121
Db 61 TELLFDAIERLGIATYFEEINQVFNHTYAYGDKWTGNTSLFRLNKGQVYVTC 120
QY 122 VFNNYKXNGAFKQSLANDVEGLLELYEATSMRVPGHIELEDALGFRSLRSLTWTDAFS 181
Db 121 IFTYKDKGRKESLEKDVHGLLELYEAYMVFVPGELIDDLAVFRTCLDRIAKNPSL 180
QY 182 TNPALFTETQALKQPLKRLPRIEAAQYIPFYQQDQSHNKTLLKAKLFNLLQSLHKE 241

```

```

Db 181 SNSAVSSQIREALTOPLHKRLPREALRYIPFYQQQASHSETLLKAKLGFNQLQSLHK 240
QY 242 ELSHVCKWKAKAFDKKNAPCLRDRIVECYFWGLSGGYEPOYSRARVPFTKAVAVITLDD 301
Db 241 ELSIISKWKSFDVANNLPYARNRVECYFWALAVYFEPQYSESRVFLSRFSTQTFDD 300
QY 302 TYDAYGTYEELKIFTEAVERWSITCLDTLPYMKPIYKLFMDTYTEWEEFLAKEQRTDLF 361
Db 301 TYDAYGTYEELEQFTRAIQWSITCLDGLPESMKLIFQMLVKIFEEIBEILSKDGKHV 360
QY 362 NCGKEFVKFVRNLMAVEKAWNEGHIPTEHHDVPIITGGANLLTTCYLGMSDIFYKE 421
Db 361 NYIKETLKEAVQSYMTAEARWAKBEYIPTIEHTKVSYSIGYKLALVAGFACMGDVIADD 420
QY 422 SVEWAVSAPPLFRYSIGILGRRLNDLMTKAEORKHSSSSLESYMKYNNVEEVAQTLLY 481
Db 421 SFWEVFNPLVACCLLCRTMDLGLSHKGEQDRKHVASTIECYMKQPDASEQAYESLN 480
QY 482 KEVEDVKDINREY-LTTKNIPRLMAVYLCQFLEVOYAGKDNFTRMGDYKHLIKSL 540
Db 481 KKVEDAWKEINREFMITCKDNIHVAMRVLPNSRSDVLYKNDHFTHVGVVEVINHIKSL 540
QY 541 LV 542
Db 541 FV 542

RESULT 10
Q9AXP5 PRELIMINARY; PRT; 573 AA.
ID Q9AXP5 PRELIMINARY; PRT; 573 AA.
AC Q9AXP5;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Sesquiterpene cyclase.
OS Artemisia annua (Sweet wormwood).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC campanulids; Asterales; Asteraceae; Asteroidae; Anthemideae;
OC Artemisia.
OX NCBI_TaxID=35608;
RN [1]
RP SEQUENCE FROM N.A.
RA Liu Y., Ye H.C., Li G.P.;
RX STRAIN=001;
RT "Cloning, characterization and E.coli expression of a cDNA encoding
RL sesquiterpene cyclase from Artemisia annua.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF304444; AAG24640.2; -
DR HSSP; Q40577; SEAT.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0005840; C:ribosome; IEA.
DR GO; GO:0016829; P:lyase activity; IEA.
DR GO; GO:0003735; P:structural constituent of ribosome; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001865; Ribosomal S2.
DR InterPro; IPR005630; Terpene synth C.
DR InterPro; IPR008949; Terpenoid synth.
DR InterPro; IPR008930; Terp_cyc_Toroid.
DR InterPro; IPR001906; Terp_synth-like.
DR Pfam; PF01397; Terpene synth; 1.
DR Pfam; PF03936; Terpene synth C; 1.
DR PROSITE; PS00962; RIBOSOMAL_S2_1; 1.
SQ SEQUENCE 573 AA; 66871 MW; 1880D258B19C039B CRC64;

Query Match 45.2%; Score 1314.5; DB 10; Length 573;
Best Local Similarity 47.0%; Pred. No. 3.3e-89;
Matches 257; Conservative 101; Mismatches 172; Indels 17; Gaps 5;

QY 10 IRPTANFPSPISWGQFLYKQVQGVQVQVNDLKEVRLKEALDIPMKH 62

```

Db 32 VRNTVFPSPISWGQFLTYTHPKKSLTEKOHVEV-----EIRKKLLNTAFSEPKQHT 86
QY 63 NLKLDEIORLIGIPYHFEREIDHALOCITYTYGDNWNG-----DRSSLWFLMRKQGYV 118
Db 87 KLHLDSVORLGVSHFEQIEBALQHVYATHGQWIGKDNLXSTSQWFLRQCGFNV 146
QY 119 TCDVFNKYDKNGAFKOSLANDVEGLELEBYATSRVPGEIILEDALGFTSRSLIMTKD 178
Db 147 SSGIFENHMDKGFKEHLSDDVQGLGLYEAAVMSVEGEKELDDALEFTKHLGNIAKD 206
QY 179 APTNPALEIQRALKQPLWKRLPRIEAAQYIPFYQOQDSHNKTLKLAKLENLQSL 238
Db 207 P-SQNASLRTKIBQALNRPLRKMRLEALHYIPYQOEASHDETLLNLAKLDYNNMLOSI 265
QY 239 HKEELSHVCKWKAFADIKKNAPCLDRDRIVECYFWGLSGYFPQYSRARVFTTKAVAVITL 298
Db 266 HKGEISEICKWKDKLDSNKLPHVDRDLVEIYFWLVGVYFPQYHRSFMFLTKTSMWLIV 325
QY 299 IDTYDAYGYEELKIIFTEAVERWSITCLDPLPYMKPIYKLFMDTYTEMEEFLEAKGRT 358
Db 326 LDDTYDNYGYEELKIIFTEAVERWSITCLDPLPYMKPIYKLFMDTYTEMEEFLEAKGRT 385
QY 359 DLFNCGEFKEFVRNLWAKWANEHGIPTTEEHDPVVIITGGANLLTTTCYLGMSDIF 418
Db 386 YQFYAKEMVYKCRNLLVEAKWLKEGVPTLEEHMSVSCVYAYAVMIANSYVGRNNV 445
QY 419 TKESVWAVSAPPLFRYSGLGRDLNLMTHKABQERKSHSSLESYMKYNNVEEYATQ 478
Db 446 TESFKWVAKFPPLVNAVATCLILRLMDLADIAGHEEERQERDHHVSSIECYENETGASEEDAVK 505
QY 479 LIYKEVEDVWKDINREYLTNKNIPRLPMVAVIYLCQFLEVQYAGKDNFTRMGDEYKHLIK 538
Db 506 FLKQVEDAKVINKELRPTPEIPMSLTIPPLSLARVSDILYKTNNGYNHAGEEIVISYIK 565
QY 539 SLAVYPM 545
Db 566 SLIVHPL 572

RESULT 11

Q9FYX7 PRELIMINARY; PRT; 577 AA.
ID Q9FYX7
AC Q9FYX7
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative sesquiterpene cyclase.
GN CASC125.
OS Artemisia annua (Sweet wormwood).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC campanulids; Asterales; Asteraceae; Asteroideae; Anthemideae;
OC Artemisia.
CX NCBI_TaxID=35608;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RX PubMed=10996256;
RA Van Geldre E.G.; De Pauw I.; Inze D.; Van Montagu M.;
RA Van den Beekhout E.;
RT "Cloning and molecular analysis of two new sesquiterpene cyclases from
RT Artemisia annua L."
RL Plant Sci. 158:163-171(2000).
DR EMBL; AJ271792; CAC12731.1; -.
DR HSP; Q4577; SEAT.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR005630; Terpene synth C.
DR InterPro; IPR008949; Terpenoid synth.
DR InterPro; IPR008930; Terp_cyc Toroid.
DR InterPro; IPR001906; Terp_synth-like.
DR Pfam; PF01397; Terpene synth; 1.
DR Pfam; PF03936; Terpene synth C; 1.

QY 5 TEEKIRPIANPPSPISWGQFLTYTHPKKSLTEKOHVEV-----EIRKKLLNTAFSEPKQHT 86
Db 26 TKPDVIRHTMNFNASIWGQFLTYDEPDLVMMKQLVEELKEEVK--KELITTKGSENP 82
QY 59 MKHANLLKLDIORLIGIPYHFEREIDHALOCITYTYGDNWNG-----NGDRSSLWFLMRKQ 114
Db 83 MOHVKLIELIDAVQRIGIAYHFEELIEBALQHVYATHGQWIGKDNLXSTSQWFLRQCGFNV 146
QY 115 GYVTCDFVFNKYDKNGAFKOSLANDVEGLELEBYATSRVPGEIILEDALGFTSRSLIMTKD 178
Db 143 GNVSSGVFKDFMDEKGFKEHLSDDVQGLGLYEAAVMSVEGEKELDDALEFTKHLGNIAKD 206
QY 175 MYKDAFSTNPALEIQRALKQPLWKRLPRIEAAQYIPFYQOQDSHNKTLKLAKLENLQSL 234
Db 203 IAKDP-SCDSSLRTQIHALKQPLRRRLARIEALHYIPYQOEASHDETLLNLAKLDYNNMLOSI 265
QY 235 LQSLHKEELSHVCKWKAFADIKKNAPCLDRDRIVECYFWGLSGYFPQYSRARVFTTKAVAVITL 298
Db 262 LQSMKELSHVCKWKAFADIKKNAPCLDRDRIVECYFWGLSGYFPQYSRARVFTTKAVAVITL 325
QY 295 VITLIDDTYDAYGYEELKIIFTEAVERWSITCLDPLPYMKPIYKLFMDTYTEMEEFLEAKGRT 358
Db 322 MLVLDLDDTYDNYGYEELKIIFTEAVERWSITCLDPLPYMKPIYKLFMDTYTEMEEFLEAKGRT 385
QY 355 BORTDLFN--CKGEFKEFVRNLWAKWANEHGIPTTEEHDPVVIITGGANLLTTTCYLGMSDIF 418
Db 382 GKG-NISNSLCQGRWOKELGSIQTLVETKAKRGVHAQPLEEYMSVSNVGTGYGLMIARS 440
QY 411 YLGMSDIFTKESVWAVSAPPLFRYSGLGRDLNLMTHKABQERKSHSSLESYMKYNNVEEYATQ 478
Db 441 YVGRGDIVTETDFKVVSSYPPPIKASCVIVRLMDDIVSHKEEERQERDHHVSSIECYENETGASEEDAVK 505
QY 471 VNEEYATQLIYKEVEDVWKDINREYLTNKNIPRLPMVAVIYLCQFLEVQYAGKDNFTRMGDEYKHLIK 538
Db 501 ASEEEACEYISRKVEDAKVINKELRPTPEIPMSLTIPPLSLARVSDILYKTNNGYNHAGEEIVISYIK 565
QY 531 DEYKHLIKSLIVYPM 547
Db 561 GDMKSYMKSFVHPMVV 577

RESULT 12

Q9FQ27 PRELIMINARY; PRT; 544 AA.
ID Q9FQ27
AC Q9FQ27
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sesquiterpene synthase 1.
GN SSTLH1.
OS Lycopersicon hirsutum.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Solanum.
CX NCBI_TaxID=62890;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GV. LA1777;
RX MEDLINE=20541624; PubMed=11090225;
RA van Der Hoeven R.S.; Monforte A.J.; Breeden D.; Tanksley S.D.;
RA Steffens J.C.;
RT "Genetic Control and Evolution of Sesquiterpene Biosynthesis in
RT Lycopersicon esculentum and Lycopersicon hirsutum."
RL Plant Cell 12:2283-2294(2000).
DR EMBL; AF279455; AAC41891.1; -.
DR HSP; Q40577; SEAT.
DR GO; GO:0016829; F:lyase activity; IEA.


```
OC eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
OX NCBI_TaxID=29729;
RN [1]
RC STRAIN=cv. Nanjing;
RA Liang W.Q., Tan X.P., Heinstein P., Yamada Y., Hashimoto T.,
RA Chen X.Y.;
RT "Expression pattern of a (+)-delta-cadinene synthase gene CAD1-A and
RT sequiterpene aldehyde biosynthesis in seedlings of Gossypium arboreum
L.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y18484; CAA77191.1; -.
DR HSSP; Q40577; SEAT.
DR GO; GO:0016829; F-lyase activity; IEA.
DR GO; GO:0008152; P-metabolism; IEA.
DR InterPro; IPR005630; Terpene synth C.
DR InterPro; IPR008949; Terpenoid synth.
DR InterPro; IPR008930; Terp cyc toroid.
DR InterPro; IPR001906; Terp synth-like.
DR Pfam; PF01397; Terpene synth; 1.
DR Pfam; PF03936; Terpene synth C; 1.
SQ SEQUENCE 555 AA; 64104 MW; 7D0A6394B133D53B CRC64;

Query Match 39.3%; Score 1143; DB 10; Length 555;
Best Local Similarity 40.9%; Pred. No. 1.8e-76;
Matches 226; Conservative 111; Mismatches 197; Indels 18; Gaps 6;

QY 3 ALTEKEPIRTPANPPSPSTWGQFLY-KQYQVQGVQIYVNDLKKEVRLKKEALDIPMKH 61
DB 15 AISEEN--RPKADPHPGWGMFIICPTDIDAATELOVEELKAQVRKMINPEVD---DS 69
QY 62 ANLLKLDEIORLGIYPYHFEREIDHALOCITYETGDNWGD-----RSSLWFRMRKOG 115
DB 70 NQKLPFDVAVQRLGVSVYHFEIEDELENIYR---DTNNNDADTDLYTALAFRLREHG 126
QY 116 YVYTCDFVNNYKQNGAFKQSLANDVEGLLELYEATSMRVPGEEIILEDALGTRSLRM 175
DB 127 FDISCDAPNFKDEAGNFKASLTSVQGLLELYEASYNRVHGEDILDEAISFTTAQLTIA 186
QY 176 TKDAFSTNPALFTFIORALKQPLWKRIPIRIEAAQVIPPYQQODSHNKTLLKLAKLENLL 235
DB 187 LP---TLHHPLEQVGHALKQSIRGLPRVEARNFISYQDLSEHNKSLLOQAKIDFNLL 243
QY 236 QSLHKEELSHYCKWKWKAFDIKKNAPCLDRIVECYFWGLSGYEPQYSRARVFFTKAVAV 295
DB 244 QLLHKEELSEICRWKDKLDFTRKLPADRDVVEGVFWINGVYFEPQYSLGRKMLTKVIAM 303
QY 296 ITLIDDDYDAYGYEELKIFTEAVERWSITCLDTLPEYMKPIYKLFMDTYTEMEBEFLAKE 355
DB 304 ASIYDDTVDSYATVDELIPYNAIERWDIKMNQLPNTWKISYKALLDVYEBEQLLANQ 363
QY 356 GRTDLFNGCKEFVKEFVRLNVAEKANEGHIPTTEHDPVVIITGGANLLTTTCVLGMS 415
DB 364 GRQYVEYAKAMIRLVQAYLLEAKWTHQNYKPTFEERFDNALPTSGYAMLAITAPVGVG 423
QY 416 DIFTKESVEMAVSAPPLFRYSIGILGRRLNDLMTKAEQERKHSSSLESYMKYKYNNEBY 475
DB 424 EVITPEFKWAASDPKIIKASTIICRFMDIDAEHFNHREDDCSAIECYKQYGVTAQE 483
QY 476 AQTLYIYREVDVWKDINREYLTNKPRLPLMAVYLCQFLVQVAGKDNFTRMGDEYKH 535
DB 484 AYNEFNKHIESKWDVNEBEFLKPTMPTVLCRSNLARVMDVLYREGDGYTHVGKAAG 543
QY 536 LIKSLLYVPMIS 547
DB 544 GITSLLIDPIQI 555

RESULT 15
Q9FTU6 PRELIMINARY; PRT; 555 AA.
AC Q9FYU6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
```

Search completed: June 3, 2004, 16:41:38
Job time : 47 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2004, 16:34:11 ; Search time 63 Seconds
(without alignments)
2462.200 Million cell updates/sec

Title: US-09-763-822A-14

Perfect score: 2906
Sequence: 1 TMTALTEKPIRPIANFPPI.....GDEYKHLIKSLLVPMISGS 549

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2891	99.5	546	3	AAy81269 Artemisia
2	2887	99.3	546	3	AAy81272 Artemisia
3	2878	99.0	546	5	AAm50093 Armopha-4
4	2855	98.2	704	3	AAy81271 Artemisia
5	1463.5	50.4	547	5	AAm50094 Armopha-4
6	1139	39.2	554	5	AAm50096 Armopha-4
7	1135	39.1	548	2	AAy27062 L. esculi
8	1123	38.6	548	2	AAy27063 L. esculi
9	1123	38.6	548	3	AAy90856 Tomato ge
10	1115.5	38.4	556	7	Ad82708 Terpenoid
11	1100	37.9	555	7	Ad82706 Terpenoid
12	1088.5	37.5	554	7	Ad82703 Terpenoid
13	1041	35.8	556	3	AAy90843 Potato ve
14	1014.5	34.9	548	3	AAy90831 Tobacco 5
15	1010.5	34.8	548	3	AAy90833 Tobacco 5
16	1010.5	34.8	548	3	AAy90834 Tobacco 5
17	1009.5	34.7	548	3	AAy90836 Tobacco 5
18	1005.5	34.6	550	2	AAw01637 Tobacco e
19	1005.5	34.6	550	5	AAm50095 Tobacco e
20	1001.5	34.5	548	3	AAy90835 Tobacco 5
21	1000.5	34.4	548	3	AAy90832 Tobacco 5
22	988.5	34.0	558	3	AAy23175 Chicory g
23	967.5	33.3	601	3	AAy90851 Castor be
24	930.5	32.0	583	3	AAy23174 Chicory g
25	915	31.5	179	3	AAy81270 Artemisia

ALIGNMENTS

RESULT 1

AAy81269

ID AAY81269 standard; protein; 546 AA.

XX

AC AAY81269;

XX

DT 19-JUN-2000 (first entry)

XX

DE Artemisia annua amorpho-4,11-diene synthase.

XX

KW Amorpha-4,11-diene synthase; amorphadiene synthase; FPP;
farnesyl pyrophosphate; cyclisation; quinghao; sweet wormwood;
annual wormwood; sweet annie; artemisin precursor; antimalarial;
sesquiterpene; flavouring; fragrance.

XX

OS Artemisia annua.

XX

FN EP982404-A1.

XX

PD 01-MAR-2000.

XX

PF 27-AUG-1998; 98EP-00202854.

XX

PR 27-AUG-1998; 98EP-00202854.

XX

(WALL/) WALLAART T E.

XX

Wallaart TED, Bouwmeester HJ;

XX

WPI; 2000-258617/23.

DR

N-PSDB; AA292754.

XX

New isolated DNA sequences and polypeptides comprising amorpho-4,11-diene

synthase activity, useful for production of amorphadiene and/or

artemisinin.

PT

Disclosure; Fig 9A; 41pp; English.

XX

This sequence represents Artemisia annua amorpho-4,11-diene synthase.

This enzyme catalyses the cyclisation of farnesyl pyrophosphate (FPP), a

ubiquitous precursor of primary and secondary metabolites, into amorpho-

4,11-diene. Amorphadiene is a precursor of artemisin, a sesquiterpene

lactone endoperoxide produced by A. annua which is a promising

antimalarial drug candidate. Amorphadiene synthase is a branch point

enzyme in the artemisin biosynthetic pathway, and the cyclisation of FPP

is therefore likely to be the rate-limiting step. Artemisin is present in

very low concentrations in A. annua, making it expensive for use as an

antimalarial drug, and attempts at organic synthesis have been

CC unsuccessful. Sesquiterpenes such as amorphadiene are also useful as
CC flavour and fragrance compounds in the food and perfume industries.
CC Terpenes also play a role in plant-insect interactions, such as the
CC attraction or repulsion of insects by plants. In addition,
CC dihydroartemisinic acid, an intermediate in the metabolic route from
CC amorphadiene to artemisinin in *A. annua*, can be used as an antioxidant.
CC DNA encoding amorphadiene synthase may be used to generate transgenic
CC plants able to synthesise amorphadiene synthase. Preferably, the plants
CC used for amorphadiene production are plants which naturally produce
CC sesquiterpenes, as these plants already have the basic sesquiterpene
CC synthetic pathways and storage compartments. *A. annua* may be transformed
CC with amorphadiene synthase expression constructs, enabling more
CC amorphadiene to be produced and hence increasing artemisinin production. It
CC may then be economically feasible to extract artemisinin from such plants
CC for use as an antimalarial
XX
XX
SQ Sequence 546 AA;

Query Match 99.5%; Score 2891; DB 3; Length 546;
Best Local Similarity 100.0%; Pred. No. 5.7e-250;
Matches 546; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 MALTEKPIRPIANPPPSIWGDQFLIYQKQVEQVQVNDLKKEVRQLLKEALDIPMKH 61
DB 1 MALTEKPIRPIANPPPSIWGDQFLIYQKQVEQVQVNDLKKEVRQLLKEALDIPMKH 60
QY 62 ANLLKLIDEIQRIGIPYHFEREIDHALQCIYETYGDNWNGDRSSLWFLMRKQGYVTCD 121
DB 61 ANLLKLIDEIQRIGIPYHFEREIDHALQCIYETYGDNWNGDRSSLWFLMRKQGYVTCD 120
QY 122 VFNNYKDKNGAFKQSLANDVEGLELLELYEATSMRVPGEIILEDALGFTSRSLIMTKDAFS 181
DB 121 VFNNYKDKNGAFKQSLANDVEGLELLELYEATSMRVPGEIILEDALGFTSRSLIMTKDAFS 180
QY 182 TNPALFTEIQRALKQPLWKRLPRIEAAQYIPFYQQQDSHNKTLILKAKLFNLLQSLHKE 241
DB 181 TNPALFTEIQRALKQPLWKRLPRIEAAQYIPFYQQQDSHNKTLILKAKLFNLLQSLHKE 240
QY 242 ELSHVCKWMAFDIKKNAFCPLDRIVECYFWGLSGVGEPOYSRARVFTKRAVAVITLIDD 301
DB 241 ELSHVCKWMAFDIKKNAFCPLDRIVECYFWGLSGVGEPOYSRARVFTKRAVAVITLIDD 300
QY 302 TYDAYGTVEELKIFTEAVERWSITCLDTPLEYMKPIYKLFMDITYTEMEEFLLAKEGRDLDL 361
DB 301 TYDAYGTVEELKIFTEAVERWSITCLDTPLEYMKPIYKLFMDITYTEMEEFLLAKEGRDLDL 360
QY 362 NCGKEFVKFVRNLMYEAKEWANEHGIPTTEHDPVVIITGGANLLTTTCYLGMSDIFTKE 421
DB 361 NCGKEFVKFVRNLMYEAKEWANEHGIPTTEHDPVVIITGGANLLTTTCYLGMSDIFTKE 420
QY 422 SYEWAASAPLFRYSIGILGRRLNDLMTKAEQERKSSSLESYMKKEYNVNVEYAQTLIY 481
DB 421 SYEWAASAPLFRYSIGILGRRLNDLMTKAEQERKSSSLESYMKKEYNVNVEYAQTLIY 480
QY 482 KEVEDVWMDINREYLTTKNIIPRLMAVIYLCQFLEVQYAGKDNFTRMGDEYKHLIKSL 541
DB 481 KEVEDVWMDINREYLTTKNIIPRLMAVIYLCQFLEVQYAGKDNFTRMGDEYKHLIKSL 540
QY 542 VYPMISI 547
DB 541 VYPMISI 546

RESULT 2
AA581272
ID AA581272 standard; protein; 546 AA.
XX
XX
XX
XX
DT 19-JUN-2000 (first entry)
XX
DE Artemisia annua amorphadiene synthase, encoded by DNA clone Z92763.
XX

KW Amorpha-4,11-diene synthase; amorphadiene synthase; FPP;
KW farnearyl pyrophosphate; cyclisation; artemisinin precursor; antimalarial;
KW sesquiterpene; flavouring; fragrance.
XX
OS Artemisia annua.
XX EP982404-A1.
XX
PD 01-MAR-2000.
XX
XX 27-AUG-1998; 98EP-00202854.
XX
XX 27-AUG-1998; 98EP-00202854.
XX
XX (WALL/) WALLAART T E.
XX
XX Wallaart TED, Bouwmeester HJ;
XX
XX WPI; 2000-258617/23.
XX N-PSDB; AAZ92763.
XX
XX New isolated DNA sequences and polypeptides comprising amorpha-4,11-diene
XX synthase activity, useful for production of amorphadiene and/or
XX artemisinin.
PS Disclosure; Page 14-16; 41pp; English.

The invention relates to DNA encoding Artemisia annua amorpha-4,11-diene
synthase (AAZ92754). This enzyme catalyses the cyclisation of farnearyl
pyrophosphate (FPP), a ubiquitous precursor of primary and secondary
metabolites, into amorpha-4,11-diene. Amorphadiene is a precursor of
artemisin, a sesquiterpene lactone endoperoxide produced by *A. annua*
which is a promising antimalarial drug candidate. Amorphadiene synthase
is a branch point enzyme in the artemisin biosynthetic pathway, and the
cyclisation of FPP is therefore likely to be the rate-limiting step.
Artemisin is present in very low concentrations in *A. annua*, making it
expensive for use as an antimalarial drug, and attempts at organic
synthesis have been unsuccessful. Sesquiterpenes such as amorphadiene are
also useful as flavour and fragrance compounds in the food and perfume
industries. Terpenes also play a role in plant-insect interactions, such
as the attraction or repulsion of insects by plants. In addition,
dihydroartemisinic acid, an intermediate in the metabolic route from
amorphadiene to artemisinin in *A. annua*, can be used as an antioxidant.
DNA encoding amorphadiene synthase may be used to generate transgenic
plants able to synthesise amorphadiene synthase. Preferably, the plants
used for amorphadiene production are plants which naturally produce
sesquiterpenes, as these plants already have the basic sesquiterpene
synthetic pathways and storage compartments. *A. annua* may be transformed
with amorphadiene synthase expression constructs, enabling more
amorphadiene to be produced and hence increasing artemisinin production. It
may then be economically feasible to extract artemisinin from such plants
for use as an antimalarial. The present sequence represents *A. annua*
amorphadiene synthase encoded by DNA clone AAZ92763; neither of these
sequences are referred to further in the specification

Sequence 546 AA;

Query Match 99.3%; Score 2887; DB 3; Length 546;
Best Local Similarity 99.6%; Pred. No. 1.3e-249;
Matches 544; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 MALTEKPIRPIANPPPSIWGDQFLIYQKQVEQVQVNDLKKEVRQLLKEALDIPMKH 61
DB 1 MSLTEKPIRPIANPPPSIWGDQFLIYQKQVEQVQVNDLKKEVRQLLKEALDIPMKH 60
QY 62 ANLLKLIDEIQRIGIPYHFEREIDHALQCIYETYGDNWNGDRSSLWFLMRKQGYVTCD 121
DB 61 ANLLKLIDEIQRIGIPYHFEREIDHALQCIYETYGDNWNGDRSSLWFLMRKQGYVTCD 120
QY 122 VFNNYKDKNGAFKQSLANDVEGLELLELYEATSMRVPGEIILEDALGFTSRSLIMTKDAFS 181
DB 121 VFNNYKDKNGAFKQSLANDVEGLELLELYEATSMRVPGEIILEDALGFTSRSLIMTKDAFS 180

QY 182 TNPALFTEIQRLKQPLWKRLPRIEAAQYIPFYQQODSHNKTLLKLAKEFNLLQSLHKE 241
DB 181 TNPALFTEIQRLKQPLWKRLPRIEAAQYIPFYQQODSHNKTLLKLAKEFNLLQSLHKE 240
QY 242 ELSHVCKWKAFDIKKNAPCLDRIVECYFWGLSGGYEPQYSRVRVFTKAVAVITLIDD 301
DB 241 ELSHVCKWKAFDIKKNAPCLDRIVECYFWGLSGGYEPQYSRVRVFTKAVAVITLIDD 300
QY 302 TYDAYGYEELKIIFTEAVERWSITCLDITLPEYMKPIYKLFMDTYTEMBEFLAKEGRTDLF 361
DB 301 TYDAYGYEELKIIFTEAVERWSITCLDITLPEYMKPIYKLFMDTYTEMBEFLAKEGRTDLF 360
QY 362 NCGKEFKVKEFVRNLMVEAKWANEHGIPTTEBHPVVIITGGANLLTTTCYLGMSDIFTKE 421
DB 361 NCGKEFKVKEFVRNLMVEAKWANEHGIPTTEBHPVVIITGGANLLTTTCYLGMSDIFTKE 420
QY 422 SVEWAVSAPPLFRYSGLGRRLNDLMTKAEQERKHSSSSLESYMKYNNVNEEYQATLIY 481
DB 421 SVEWAVSAPPLFRYSGLGRRLNDLMTKAEQERKHSSSSLESYMKYNNVNEEYQATLIY 480
QY 482 KEVEDVWKDINREYLTQNIIPRLMAVIYLCQFLEVQYAGKDNFTRMGDEYKHLIKSLL 541
DB 481 KEVEDVWKDINREYLTQNIIPRLMAVIYLCQFLEVQYAGKDNFTRMGDEYKHLIKSLL 540
QY 542 VYPMSI 547
DB 541 VYPMSI 546

RESULT 3
ID AAM50093 standard; protein; 546 AA.
AC AAM50093;
XX
DT 20-AUG-2002 (first entry)
XX
DE Armopha-4,11-diene synthase associated protein #1.
DE
KW Armopha-4,11-diene synthase; armopha-4,11-diene; antimalaria agent;
KW artemisinin biosynthesis.
XX
OS Artemisia annua.
XX
PN KR2001084864-A.
XX
PD 06-SEP-2001.
XX
PF 29-FEB-2000; 2000KR-00010207.
XX
PR 29-FEB-2000; 2000KR-00010207.
XX
PA (KIMS/) KIM S U.
XX
PI Jang YJ, Kim SU, Park SH, Song SH;
XX
PI WPI; 2002-153413/20.
XX
XX
PT Armopha-4,11-diene synthase of artemisia annua, useful for mediating
PT synthesis of antimalaria agent armopha-4,11-diene.
XX
PS Disclosure; Fig 7; 17pp; Korean.
XX
CC This invention describes a novel armopha-4,11-diene synthase of Artemisia
CC annua, which mediates the synthesis of armopha-4,11-diene in the
CC biosynthesis pathway of artemisinin and is useful as an antimalaria
CC agent. This sequence represents a protein associated with the invention
XX
SQ Sequence 546 AA;

Query Match 99.0%; Score 2878; DB 5; Length 546;
Best Local Similarity 99.3%; Pred. No. 8.4e-249;
Matches 542; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 MALTTEKPIRPIANFPSSINGDOFLIYKQVQGVQEQIVNDLKKEVROLLKEALDIPMKH 61
DB 1 MSLTEKPIRPIANFPSSINGDOFLIYKQVQGVQEQIVNDLKKEVROLLKEALDIPMKH 60
QY 62 ANLLKLIIDEIQRLGIPYHFEREIDHALQCIYETYGDNWNGDRSSLWFLMRKQGYVTC 121
DB 61 ANLLKLIIDEIQRLGIPYHFEREIDHALQCIYETYGDNWNGDRSSLWFLMRKQGYVTC 120
QY 122 VFNNYKDKGAFKQSLANDVEGLELLEYEATSMRVPGBIILEDALGFTSRSLSMNTKDAPS 181
DB 121 VFNNYKDKGAFKQSLANDVEGLELLEYEATSMRVPGBIILEDALGFTSRSLSMNTKDAPS 180
QY 182 TNPALFTEIQRLKQPLWKRLPRIEAAQYIPFYQQODSHNKTLLKLAKEFNLLQSLHKE 241
DB 181 TNPALFTEIQRLKQPLWKRLPRIEAAQYIPFYQQODSHNKTLLKLAKEFNLLQSLHKE 240
QY 242 ELSHVCKWKAFDIKKNAPCLDRIVECYFWGLSGGYEPQYSRVRVFTKAVAVITLIDD 301
DB 241 ELSHVCKWKAFDIKKNAPCLDRIVECYFWGLSGGYEPQYSRVRVFTKAVAVITLIDD 300
QY 302 TYDAYGYEELKIIFTEAVERWSITCLDITLPEYMKPIYKLFMDTYTEMBEFLAKEGRTDLF 361
DB 301 TYDAYGYEELKIIFTEAVERWSITCLDITLPEYMKPIYKLFMDTYTEMBEFLAKEGRTDLF 360
QY 362 NCGKEFKVKEFVRNLMVEAKWANEHGIPTTEBHPVVIITGGANLLTTTCYLGMSDIFTKE 421
DB 361 NCGKEFKVKEFVRNLMVEAKWANEHGIPTTEBHPVVIITGGANLLTTTCYLGMSDIFTKE 420
QY 422 SVEWAVSAPPLFRYSGLGRRLNDLMTKAEQERKHSSSSLESYMKYNNVNEEYQATLIY 481
DB 421 SVEWAVSAPPLFRYSGLGRRLNDLMTKAEQERKHSSSSLESYMKYNNVNEEYQATLIY 480
QY 482 KEVEDVWKDINREYLTQNIIPRLMAVIYLCQFLEVQYAGKDNFTRMGDEYKHLIKSLL 541
DB 481 KEVEDVWKDINREYLTQNIIPRLMAVIYLCQFLEVQYAGKDNFTRMGDEYKHLIKSLL 540
QY 542 VYPMSI 547
DB 541 VYPMSI 546

RESULT 4
AAY81271
ID AAY81271 standard; protein; 704 AA.
XX
AC AAY81271;
XX
DT 19-JUN-2000 (first entry)
XX
DE Artemisia annua amorphadiene synthase, encoded by a positive DNA clone.
XX
KW Amopha-4,11-diene synthase; amorphadiene synthase; PPP;
KW farnesyl pyrophosphate; cyclisation; artemisin precursor; antimalarial;
KW sesquiterpene; flavouring; fragrance.
XX
OS Artemisia annua.
XX
FH Key Location/Qualifiers
FT Misc-difference 267 /note= "Encoded by CCT"
FT Misc-difference 270 /note= "Encoded by AGA"
FT Misc-difference 272 /note= "Encoded by AGA"
FT Misc-difference 291 /note= "Encoded by AGA"
FT Misc-difference 555 /label= Xaa
FT /note= "Encoded by TGA, Xaa = unknown"
FT Misc-difference 562 /label= Xaa
FT /note= "Encoded by TAG, Xaa = unknown"

Db 418 FISHTEFFELMNESVIVRASALIARANDIVGHEDQERGHVSLIECYKWDYGASKQET 477
 Qy 477 QFLIYKEVDVWKNINREYLTTRNIPRLPMVILYCOFLEVYAGKDNFRMGDEYKHL 536
 Db 478 YIKFLKEVTNAWKDINKQSPREVEWMFVLNLRVADILYKEDTYSYTAGKLNKM 537
 Qy 537 IKSLIYVPMISI 547
 Db 538 INPILESVKI 548

RESULT 10

ADE82708

ID ADE82708 standard; protein; 556 AA.

XX AC ADE82708;

XX DT 29-JAN-2004 (first entry)

XX DE Terpenoid biosynthesis related H64 strawberry protein #14.

XX KW isoprenoid; bio-active compound synthesis; pesticide; dermatological;

XX KW cytostatic; immunosuppressive; virucide; flavour; fragrance;

XX KW bio-control agent; food additive; food industry; pest control;

XX KW degreasing solvent; plasticizer; dye carrier; dental caries;

XX KW dental plaque; skin disorder; immunosuppressive; anti-leukaemia;

XX KW anti-retroviral; monoterpene alcohol linalool;

XX KW sesquiterpene alcohol nerolidol; monoterpene; strawberry.

XX OS Fragaria x ananassa.

XX PN EP1231273-A1.

XX PD 14-AUG-2002.

XX PF 12-FEB-2001; 2001EP-00200488.

XX PR 12-FEB-2001; 2001EP-00200488.

XX PA (PLAN-) PLANT RES INT BV.

XX PI Aharoni A, Verhoeven HA, Jongema MA, Bouwmeester HJ;

XX DR WPI; 2003-879727/82.

XX PT Novel recombinant nucleic acid encoding proteinaceous molecule, useful

XX PT for producing flavor, fragrance and/or biocontrol agent which is useful

XX PT as food additive in processed food industry and as antimicrobial agent.

XX PS Disclosure; Page; 52pp; English.

XX CC The invention relates to a novel isolated or recombinant nucleic acid or

XX CC its functional fragment, encoding a proteinaceous molecule essentially

XX CC capable of isoprenoid bio-active compound synthesis when provided with a

XX CC suitable substrate under appropriate reaction conditions. The novel

XX CC recombinant isoprenoid bio-active compound synthesis nucleic acid and its

XX CC protein have the following activities: pesticide, dermatological,

XX CC cytostatic, immunosuppressive, and virucide. The novel recombinant

XX CC isoprenoid bio-active compound synthesis nucleic acid is useful for

XX CC producing flavour, fragrance, and/or a bio-control agent, by transforming

XX CC or transfecting a suitable host with the recombinant isoprenoid bio-

XX CC active compound synthesis nucleic acid, expressing the recombinant

XX CC isoprenoid bio-active compound synthesis nucleic acid in the presence of

XX CC a suitable substrate, and optionally isolating the formed product. The

XX CC bio-control agent is useful as an anti-microbial agent, as a food

XX CC additive in the processed food industry to modify the taste of syrups,

XX CC ice-creams, frozen desserts, yogurts, confectionary and like products, as

XX CC a flavouring agent for oral medications and vitamins, and for providing

XX CC additional flavour/aroma in beverages, including alcoholic beverages. The

XX CC bio-control agent is also useful for enhancing or reducing flavour,

XX CC aroma, fragrance or scent of plants, natural products, and/or synthetic

XX CC or artificial products, and for the industrial synthesis of nature

CC identical flavour/aroma substances and/or artificial flavour/aroma
 CC substances. The bio-control agent is also useful as a pest control agent
 CC for the biological control of the interaction between plants and insects
 CC and/or plants and microorganisms, for providing flavour/aroma in
 CC cosmetics, creams, sun-protectant products, hair conditioners, cleaning
 CC products, personal care products and health care products, as a
 CC disinfectant additive and in the preparation of a composition. The novel
 CC recombinant isoprenoid bio-active compound synthesis nucleic acid or its
 CC fragments is useful as a molecular marker or diagnostic tool. The protein
 CC of the novel recombinant isoprenoid bio-active compound synthesis nucleic
 CC acid is useful for the production of an antagonist e.g. an antibody or
 CC its functional equivalent which is useful for inhibiting the synthesis of
 CC the bio-control agent. A composition, containing the bio-control agent,
 CC is a pharmaceutical or nutraceutical, useful for augmenting or enhancing
 CC the aroma and/or taste of food or non-food products, and/or protection of
 CC food or non-food products against fungal contamination and/or pest
 CC infestation. The composition is also useful for the biological control of
 CC pests, for the protection of stored products and for the prevention or
 CC treatment of disease. The bio-control agent is useful as a degreasing
 CC solvent, plasticizer and dye carrier. The composition is useful for
 CC replacing potentially carcinogenic synthetic food additives currently
 CC used. The composition is also useful for treating dental caries, dental
 CC plaque and skin disorders, and for immunosuppressive, anti-leukaemia and
 CC anti-retroviral treatment. The novel recombinant isoprenoid bio-active
 CC compound synthesis nucleic acid or its protein is useful for the
 CC synthesis of monoterpene alcohol linalool and sesquiterpene alcohol
 CC nerolidol, and monoterpene. This sequence represents an H64 protein
 CC used in the terpenoid biosynthesis method of the invention. NOTE: This
 CC sequence is not shown in the specification. It has been obtained from
 CC electronic data supplied with this specification from the European Patent
 CC Office.

XX SQ Sequence 556 AA;

Query Match 38.4%; Score 1115.5; DB 7; Length 556;

Best Local Similarity 40.9%; Pred. NO. 9.1e-91;

Matches 226; Conservative 113; Mismatches 195; Indels 19; Gaps 5;

Qy 4 LTEEKPPIPIANFPSPISWGDQF-----LIYQKQVEQVEQIVNLDKKEVRQLLKALDI 57

Db 14 ISKPEVVRTANFANFPSPISWGDQF-----LIYQKQVEQVEQIVNLDKKEVRQLLKALDI 69

Qy 58 PMKHANLLKLIDEIQRLGIPYHFEREIDHALQCIYETVYCDNWNWD---RSSLWFLMRKQ 114

Db 70 --DSHQKLKLIDEIQRLGVAHYFSEIDQALERTHETQDIDHGGDLNVALRFRLLRRH 127

Qy 115 GYVYTCDFVNNYKDKNGAFKQSLANDVEGLLELYEATSNRVSGEILIEDALGFTRRRLSI 174

Db 128 GYNVSCDVNFNFKDNTNGDYKKSLSVTDLSGMLSFYEAHLRVHGEKLEELVFTTHL-- 185

Qy 175 MTKDAFTNPALFTETIQPALKQPLKRLPRIEAAQYIPYQQQDSHNKTLKLAKLEFNL 234

Db 186 --QSASAKSSLLKQTQITEAVERPLLTWELGARRYSIYQDEASYSNLLKLAKLDFNV 243

Qy 235 LQSLHKELSHVCKWKAFAIDKKNAPCLDRIVECYFWGLSGSYEPQYGRARVFTTKAVA 294

Db 244 VQCHKELSLILRWYKELDFARRPFPARRIVELFFWTAGIYFPEYVFGHILAKLIE 303

Qy 295 VTIIDDDYDAGYVEELKITEAVERMSITCLDTLPYMKPIYKLFMTYTYMESEFLAK 354

Db 304 ITTVMDMDYDAFGTPEELVILTEAIDRWDAASCMDQLPDYMQPFYITLLDVIDVEBELTK 363

Qy 355 EGRDLPNCGKEFVKEFVRNLMVKAQNEGHIPTEEHDPVVIITGGANLLTTCVLGM 414

Db 364 QGRSVRIHYAKIEKMNQARLYFAEAIAWPHFEGCTPKMDGYNRVAASVGNLTSLVSLVGM 423

Qy 415 SDIETKESVWAVSAPPLFRYSGLILGRRLINDLMTKAEQERKSHSSSLESYMKYNNVEE 474

Db 424 GDITKPEFELTNEPKILRASNTIFRLMDDDIAGYKFEKERGHVASSIDCYWNEYGVSEQ 483

Qy 475 YAQTLIYKEVDVWKNINREYLTTRNIPRLPMVILYCOFLEVYAGKDNFRMGDEYK 534

Db 484 ETIDIFNKRIVDWMDKINEEFLRPTAAPVPLNRLNLRVVDLLYKRGDAFTHVGLMK 543

QY 535 HLIKSLLYPMSI 547
Db 544 DCIAAFIDPVPL 556

RESULT 11

AD82706
ID ADE82706 standard; protein; 555 AA.

XX AC ADE82706;

XX 29-JAN-2004 (first entry)

DE Terpenoid biosynthesis related H64 strawberry protein #13.

XX isoprenoid; bio-active compound synthesis; pesticide; dermatological;
KW cytosolic; immunosuppressive; virucide; flavour; fragrance;
KW bio-control agent; food additive; food industry; pest control;
KW degreasing solvent; plasticizer; dye carrier; dental caries;
KW dental plaque; skin disorder; immunosuppressive; anti-leukaemia;
KW anti-retroviral; monoterpene alcohol linalool;
KW sesquiterpene alcohol nerolidol; monoterpene; strawberry.

XX Fragaria x ananassa.

XX EP1231273-A1.

XX 14-AUG-2002.

XX 12-FEB-2001; 2001EP-00200488.

XX 12-FEB-2001; 2001EP-00200488.

XX (PLAN-) PLANT RES INT BV.

XX Aharoni A, Verhoeven HA, Jongsma MA, Bouwmeester HU;

XX WFI; 2003-879727/82.

PT Novel recombinant nucleic acid encoding proteinaceous molecule, useful
PT for producing flavor, fragrance and/or biocontrol agent which is useful
PT as food additive in processed food industry and as antimicrobial agent.

PS Disclosure; Page: 52pp; English.

CC The invention relates to a novel isolated or recombinant nucleic acid or
CC its functional fragment, encoding a proteinaceous molecule essentially
CC capable of isoprenoid bio-active compound synthesis when provided with a
CC suitable substrate under appropriate reaction conditions. The novel
CC recombinant isoprenoid bio-active compound synthesis nucleic acid and its
CC protein have the following activities: pesticide, dermatological,
CC cytostatic, immunosuppressive, and virucide. The novel recombinant
CC isoprenoid bio-active compound synthesis nucleic acid is useful for
CC producing flavour, fragrance, and/or a bio-control agent, by transforming
CC or transfecting a suitable host with the recombinant isoprenoid bio-
CC active compound synthesis nucleic acid, expressing the recombinant
CC isoprenoid bio-active compound synthesis nucleic acid in the presence of
CC a suitable substrate, and optionally isolating the formed product. The
CC bio-control agent is useful as an anti-microbial agent, as a food
CC additive in the processed food industry to modify the taste of syrups,
CC ice-creams, frozen desserts, yogurts, confectionery and like products, as
CC a flavouring agent for oral medications and vitamins, and for providing
CC additional flavour/aroma in beverages, including alcoholic beverages. The
CC bio-control agent is also useful for enhancing or reducing flavour,
CC aroma, fragrance or scent of plants, natural products, and/or synthetic
CC or artificial products, and for the industrial synthesis of nature
CC identical flavour/aroma substances and/or artificial flavour/aroma
CC substances. The bio-control agent is also useful as a pest control agent
CC for the biological control of the interaction between plants and insects
CC and/or plants and microorganisms, for providing flavour/aroma in
CC cosmetics, creams, sun-protectant products, hair conditioners, cleaning
CC products, personal care products and health care products, as a

CC disinfectant additive and in the preparation of a composition. The novel
CC recombinant isoprenoid bio-active compound synthesis nucleic acid or its
CC fragments is useful as a molecular marker or diagnostic tool. The protein
CC of the novel recombinant isoprenoid bio-active compound synthesis nucleic
CC acid is useful for the production of an antagonist e.g. an antibody or
CC its functional equivalent which is useful for inhibiting the synthesis of
CC the bio-control agent. A composition, containing the bio-control agent,
CC is a pharmaceutical or nutraceutical, useful for augmenting or enhancing
CC the aroma and/or taste of food or non-food products, and/or protection of
CC food or non-food products against fungal contamination and/or pest
CC infestation. The composition is also useful for the biological control of
CC pests, for the protection of stored products and for the prevention or
CC treatment of disease. The bio-control agent is useful as a degreasing
CC solvent, plasticizer and dye carrier. The composition is useful for
CC replacing potentially carcinogenic synthetic food additives currently
CC used. The composition is also useful for treating dental caries, dental
CC plaque and skin disorders, and for immunosuppressive, anti-leukaemia and
CC anti-retroviral treatment. The novel recombinant isoprenoid bio-active
CC compound synthesis nucleic acid or its protein is useful for the
CC synthesis of monoterpene alcohol linalool and sesquiterpene alcohol
CC nerolidol, and monoterpene. This sequence represents an H64 protein
CC used in the terpenoid biosynthesis method of the invention. NOTE: This
CC sequence is not shown in the specification. It has been obtained from
CC electronic data supplied with this specification from the European Patent
CC Office.

XX

SQ Sequence 555 AA;

Query Match 37.9%; Score 1100; DB 7; Length 555;

Best Local Similarity 41.1%; Pred. No. 2.2e-89; Indels 20; Gaps 6;
Matches 225; Conservative 110; Mismatches 192;

QY 10 IRPIANPPPSINGDQF-----LIYQKQVQGVQVIVNDLKKVRLQLKEALDIPMKHAN 63
Db 20 VRTANFKPSVGMGRFANYAEDIIITQMQBQVEELAQVVRKEV---FTNAAD---DSSH 73
QY 64 LKLDIEIORLGIPIYFEBIDHALQCIYETIYDGNWGD---RSLFRLMRKQGYVTC 120
Db 74 QKLDIEIORLGVAYHFESBIDQALERIETHETQDIHDGDLVNVALRFLRRHGYNVSC 133
QY 121 DVFNKYKNGKAFKQSLANDVEGLLELYEATSRVPGEIILEDALGTRSLRSLMTKDAF 180
Db 134 DVFNKFTNGDYKSLVTDLSGMLSFYEAHARVHGEKULEEALVFTTHL-----QAS 189
QY 181 STNPALFTEIORALKQPLMKRLPRIEAAQYIPFYQQQDSHNKTLKLAKLEFNLSLHK 240
Db 190 AKSSLLKTQITEAVER-LLKTMERLGARRYMSIYQDEASYSNLLKLAKLDFNVVQLHK 248
QY 241 BELSHVCKWKAFFDIKKNAPCLDRIVECYFWGLSGGYEPOYSRPFVFFKAVAVITLID 300
Db 249 KELSILRWYKELDFARRMPFARDRIVELFFWTAGIYFEPEYFRRHILTKLEITTVMD 308
QY 301 DTYDAYGTVEELKI FTEAVERWSITCLDTLPEYWKPIYKLFMDTYTEMEEFLEKEGRDUL 360
Db 309 DMVDFAGTVEELVNLTEADRDWDASCDQLPDTWQPFYITLLDVIDVEBELTKQGSYR 368
QY 361 FNCGKEFVKFVRNLMVEAKWANEHGIPTTEEDHPVVIITGGANLLTTCYVLGMSDIFTK 420
Db 369 IHVAKELMKNQARLYPAEARWPHGCTPKWDEYMRVAASSVGNMTLSVWSLVGMGDIITK 428
QY 421 ESVEHVAVSAPPLFRVSGILGRRLNDLMTKAEORKHSSSSLESYMKYNNVNEEYAQTLLI 480
Db 429 FEFELTNPEFKILRASNTIFRLMDDIAGYKFEKRGHVASSIDCYNNYGVSEQETIDIF 488
QY 481 YKEVEDVMKDINREYLTNNIPRLPLMAVILYLCQLEVOYAGKDNFRMGDEYKHLKSL 540
Db 489 NGRIVDSWKDINEEFURPTAAPVPLNRLNLRVVDLTKRGDAFTHVGLMKDKCIAAM 548
QY 541 LVIYPMISI 547
Db 549 FIDPVPL 555

Search completed: June 3, 2004, 16:40:46
Job time : 71 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2004, 16:30:26 ; Search time 23 Seconds
(without alignments)
1232.289 Million cell updates/sec

Title: US-09-763-822A-14
Perfect score: 2906
Sequence: 1 TVALTEKPIRPIANFPPI.....GDEYKHLIKSLVPMSTGS 549

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pap:*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pap:*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pap:*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pap:*
5: /cgn2_6/ptodata/2/iaa/PTCUS COMB.pap:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1135	39.1	548	US-09-601-091-2	Sequence 2, Appli
2	1123	38.6	548	US-09-601-091-4	Sequence 4, Appli
3	1123	38.6	548	US-09-398-395A-52	Sequence 52, Appl
4	1123	38.6	548	US-09-887-586A-52	Sequence 52, Appl
5	1123	38.6	548	US-09-903-012B-52	Sequence 52, Appl
6	1123	38.6	548	US-09-900-797-52	Sequence 52, Appl
7	1123	38.6	548	US-09-900-797-52	Sequence 52, Appl
8	1041	35.8	556	US-09-398-395A-32	Sequence 32, Appl
9	1041	35.8	556	US-09-887-586A-32	Sequence 32, Appl
10	1041	35.8	556	US-09-895-752-32	Sequence 32, Appl
11	1041	35.8	556	US-09-903-012B-32	Sequence 32, Appl
12	1041	35.8	556	US-09-900-797-32	Sequence 32, Appl
13	1020.5	35.1	548	US-09-398-395A-2	Sequence 2, Appli
14	1020.5	35.1	548	US-09-887-586A-2	Sequence 2, Appli
15	1020.5	35.1	548	US-09-895-752-2	Sequence 2, Appli
16	1020.5	35.1	548	US-09-903-012B-2	Sequence 2, Appli
17	1020.5	35.1	548	US-09-900-797-2	Sequence 2, Appli
18	1017.5	35.0	548	US-09-398-395A-8	Sequence 8, Appli
19	1017.5	35.0	548	US-09-887-586A-8	Sequence 8, Appli
20	1017.5	35.0	548	US-09-895-752-8	Sequence 8, Appli
21	1017.5	35.0	548	US-09-903-012B-8	Sequence 8, Appli
22	1017.5	35.0	548	US-09-900-797-8	Sequence 8, Appli
23	1010.5	34.8	548	US-09-398-395A-6	Sequence 6, Appli
24	1010.5	34.8	548	US-09-887-586A-6	Sequence 6, Appli
25	1010.5	34.8	548	US-09-895-752-6	Sequence 6, Appli
26	1010.5	34.8	548	US-09-903-012B-6	Sequence 6, Appli
27	1010.5	34.8	548	US-09-900-797-6	Sequence 6, Appli

28	1009.5	34.7	548	US-09-398-395A-12	Sequence 12, Appl
29	1009.5	34.7	548	US-09-887-586A-12	Sequence 12, Appl
30	1009.5	34.7	548	US-09-895-752-12	Sequence 12, Appl
31	1009.5	34.7	548	US-09-903-012B-12	Sequence 12, Appl
32	1009.5	34.7	548	US-09-900-797-12	Sequence 12, Appl
33	1005.5	34.6	550	US-08-443-639-8	Sequence 8, Appli
34	1005.5	34.6	550	US-08-577-483-8	Sequence 8, Appli
35	1005.5	34.6	550	US-09-435-380-8	Sequence 8, Appli
36	1001.5	34.5	548	US-09-398-395A-10	Sequence 10, Appl
37	1001.5	34.5	548	US-09-887-586A-10	Sequence 10, Appl
38	1001.5	34.5	548	US-09-895-752-10	Sequence 10, Appl
39	1001.5	34.5	548	US-09-903-012B-10	Sequence 10, Appl
40	1001.5	34.5	548	US-09-900-797-10	Sequence 10, Appl
41	1000.5	34.4	548	US-09-398-395A-4	Sequence 4, Appli
42	1000.5	34.4	548	US-09-887-586A-4	Sequence 4, Appli
43	1000.5	34.4	548	US-09-895-752-4	Sequence 4, Appli
44	1000.5	34.4	548	US-09-903-012B-4	Sequence 4, Appli
45	1000.5	34.4	548	US-09-900-797-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-09-601-091-2
; Sequence 2, Application US/09601091
; Patent No. 6342380
; GENERAL INFORMATION:
; APPLICANT: Colby, S. M. et al.
; TITLE OF INVENTION: Germacrene C Synthase Gene of Lycopersicon Esculentum
; FILE REFERENCE: 4630-55758
; CURRENT APPLICATION NUMBER: US/09/601,091
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: PCT/US99/02133
; PRIOR FILING DATE: 1999-02-02
; PRIOR APPLICATION NUMBER: US 60/073, 579
; PRIOR FILING DATE: 1998-02-02
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 2
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Lycopersicon esculentum
US-09-601-091-2

Query Match 39.1%; Score 1135; DB 4; Length 548;
Best Local Similarity 39.2%; Pred. No. 1.3e+02;
Matches 216; Conservative 122; Mismatches 205; Indels 8; Gaps 3;
QY 2 MALTTEKPIRPIANPPPSIWGQFLYQXQVQGVQIYNDLKKVQQLLKEALDIPMKH 61
Db 1 MAASADKCRPLANFHPSPVWGYPFLSYTHEITNQEKVEVDEYKETIRKMLVETCD--NS 57
QY 62 ANLLXLIIDEIORLIGIPYHFEREIDHALQCIYETYGDNWNGDRS-----SLAFRLMRKQGY 117
Db 58 TOKVLIDAMQRLGVAIHFDNEIETSIQIFDASSKQNDNDNDVYVSLRFLRQOQHY 117
QY 118 VTCDVFNNYKDKNGAFKQSLANDVEGLELELYEATSMRVPGCEIILEDALGFTSRSLRSLMTK 177
Db 118 MSSDVFQKFTNQDGRKFKELTNDVQGLLSIYEASHLRVRNEELEBALTTTTHLESIVS 177
QY 178 DAFSTNPALFTBIQALKQPLWKRLPRIEAAQVIPPYQOODSHNKLKLLAKLEFNLLQS 237
Db 178 NLSNNNSLKVGEALTOFIRMTLPRMGARKYISYENNDAAHHLLKTKAKLDFNMLQK 237
QY 238 LHKELSHVCKWKAFDIKKNAPCLRDRIVECVFVGLSGYEPQYSPARVFFTKAVAVIT 297
Db 238 FQRELSLDRWKKOLDPANKYPADRLVECVFVGLSGYEPQYSPARVFFTKAVAVIT 297
QY 298 LIIDYDAYGTVEELKIFTEAVERWSITCLDTLPYKPKLYKLFMDTYTEMESEFLAXEGR 357
Db 298 LIIDYDAYGTVEELKIFTEAVERWSITCLDTLPYKPKLYKLFMDTYTEMESEFLAXEGR 357

QY 358 TDLFNGGKFEVKEFVRNLMVEAKWANE-GHPTTEHDPVVIITGGANLLTTTCYLGMSD 416
Db 358 LDRVYAKNEMKLVRAVYFKETQWLNDCHDHPKYEEQVENAIVSAGYMMISTTCLVGIEE 417
QY 417 IPTKESVEMAVGAPPLFRYSIGILGRINDLMTKAEQERKSHSSLESYMKYNNVEEYA 476
Db 418 FISHTEFWLWMSVIVRASALIARAMNDIVGHEDEQERGHVASLIECYMKDYGASKQET 477
QY 477 QTLIYKEVDVWKDINREYLTNIPRLLMAVIYLCQFLEVQYAGKDNFTRMGDEYKHL 536
Db 478 YIKFLKEVTNAWKDINKQFRTPEVPMFVLRLNTRVADTLYKEKDTYNAKGKLNKM 537
QY 537 IKSLIVYPMI 547
Db 538 INSILIESVKI 548

RESULT 2

US-09-601-091-4
; Sequence 4, Application US/09601091
; Patent No. 6342380
; GENERAL INFORMATION:
; APPLICANT: Colby, S. M. et al.
; TITLE OF INVENTION: Germacrine C Synthase Gene of Lycopersicon Esculentum
; FILE REFERENCE: 4630-55758
; CURRENT APPLICATION NUMBER: US/09/601,091
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: PCT/US99/02133
; PRIOR FILING DATE: 1999-02-02
; PRIOR APPLICATION NUMBER: US 60/073,579
; PRIOR FILING DATE: 1998-02-02
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Lycopersicon esculentum
US-09-601-091-4

Query Match 38.6%; Score 1123; DB 4; Length 548;
Best Local Similarity 38.8%; Pred. No. 1.9e-101;
Matches 214; Conservative 123; Mismatches 206; Indels 8; Gaps 3;
QY 2 MALTEEKPIRPIANPPPSIWGDOFLIYQKQVGEQIIVNDLKKEVRQLLKEALDIPMKH 61
Db 1 MAASADKCRPLANPHSPVWGYPHFSYTHEITNQEKVEYDEYKETIRKMLVETCD---NS 57
QY 62 ANLLKLIDEIQRIGIPYHFEREIDHALQCIYETYGDNWNGDRS----SLWFLMRKQGY 117
Db 58 TQKLVLDAMQRLGVAYHFDNEIETSIQNFDAKQNDNDNNLYVVSRLFRVLRQGHY 117
QY 118 VTCDVFNKYKDKNGAPKQSLANDVEGLLELYEATSMRVPGIIELEDALGFRSRLSIMTK 177
Db 118 MSSDVFKQFTNQDGKFKETLTNDVQGLLSYBASHLRVRNEEILEEALFTTTTHLESIVS 177
QY 178 DAFSTNPALFTIQRALKQPLKRLPRIEAAQIIPYQOQDSHNKTLKLAKLEFNLLQS 237
Db 178 NLSNNNSLKVEGEALTQPIRMTPRGMARKYISYENNDAAHLLHLLKFAKLDNFNLQK 237
QY 238 LHKEELSHVCKWKAFDIKKNAPCLDRIVECYFWGLSGSYEPQYSRARVFTTKAVAVIT 297
Db 238 FHQRELSDLTRWVKOLDPANKYPYARDLVECYFWGLGVFEPKYSRAKMMTKVNLNLS 297
QY 298 LIDDTDAYGTYEELKIFTEAVERNSITCLDTPLEYMKPIYKLFMDTYTEMEEFLLAKEGR 357
Db 298 IIDDTFDAYATFDELVTFAIQRDANALDSIQPYMRPAYQALLDIYSEMEQVLSKEGK 357
QY 358 TDLFNGGKFEVKEFVRNLMVEAKWANE-GHPTTEHDPVVIITGGANLLTTTCYLGMSD 416
Db 358 LDRVYAKNEMKLVRAVYFKETQWLNDCHDHPKYEEQVENAIVSAGYMMISTTCLVGIEE 417
QY 417 IPTKESVEMAVGAPPLFRYSIGILGRINDLMTKAEQERKSHSSLESYMKYNNVEEYA 476

Db 418 FISHTEFWLWMSVIVRASALIARAMNDIVGHEDEQERGHVASLIECYMKDYGASKQET 477
QY 477 QTLIYKEVDVWKDINREYLTNIPRLLMAVIYLCQFLEVQYAGKDNFTRMGDEYKHL 536
Db 478 YIKFLKEVTNAWKDINKQFRTPEVPMFVLRLNTRVADTLYKEKDTYNAKGKLNKM 537
QY 537 IKSLIVYPMI 547
Db 538 INSILIESVKI 548

RESULT 3

US-09-398-395A-52
; Sequence 52, Application US/09398395A
; Patent No. 6468772
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 6468772, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/398,395A
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 52
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Lycopersicon esculentum
US-09-398-395A-52

Query Match 38.6%; Score 1123; DB 4; Length 548;
Best Local Similarity 38.8%; Pred. No. 1.9e-101;
Matches 214; Conservative 123; Mismatches 206; Indels 8; Gaps 3;
QY 2 MALTEEKPIRPIANPPPSIWGDOFLIYQKQVGEQIIVNDLKKEVRQLLKEALDIPMKH 61
Db 1 MAASADKCRPLANPHSPVWGYPHFSYTHEITNQEKVEYDEYKETIRKMLVETCD---NS 57
QY 62 ANLLKLIDEIQRIGIPYHFEREIDHALQCIYETYGDNWNGDRS----SLWFLMRKQGY 117
Db 58 TQKLVLDAMQRLGVAYHFDNEIETSIQNFDAKQNDNDNNLYVVSRLFRVLRQGHY 117
QY 118 VTCDVFNKYKDKNGAPKQSLANDVEGLLELYEATSMRVPGIIELEDALGFRSRLSIMTK 177
Db 118 MSSDVFKQFTNQDGKFKETLTNDVQGLLSYBASHLRVRNEEILEEALFTTTTHLESIVS 177
QY 178 DAFSTNPALFTIQRALKQPLKRLPRIEAAQIIPYQOQDSHNKTLKLAKLEFNLLQS 237
Db 178 NLSNNNSLKVEGEALTQPIRMTPRGMARKYISYENNDAAHLLHLLKFAKLDNFNLQK 237
QY 238 LHKEELSHVCKWKAFDIKKNAPCLDRIVECYFWGLSGSYEPQYSRARVFTTKAVAVIT 297
Db 238 FHQRELSDLTRWVKOLDPANKYPYARDLVECYFWGLGVFEPKYSRAKMMTKVNLNLS 297
QY 298 LIDDTDAYGTYEELKIFTEAVERNSITCLDTPLEYMKPIYKLFMDTYTEMEEFLLAKEGR 357
Db 298 IIDDTFDAYATFDELVTFAIQRDANALDSIQPYMRPAYQALLDIYSEMEQVLSKEGK 357
QY 358 TDLFNGGKFEVKEFVRNLMVEAKWANE-GHPTTEHDPVVIITGGANLLTTTCYLGMSD 416
Db 358 LDRVYAKNEMKLVRAVYFKETQWLNDCHDHPKYEEQVENAIVSAGYMMISTTCLVGIEE 417
QY 417 IPTKESVEMAVGAPPLFRYSIGILGRINDLMTKAEQERKSHSSLESYMKYNNVEEYA 476
Db 418 FISHTEFWLWMSVIVRASALIARAMNDIVGHEDEQERGHVASLIECYMKDYGASKQET 477

QY 477 QTLIYKEVDVWKDINREYLTTKNIPRPLLMVAVIYLCQFLEVQYAGKDNFTRMGDEYKHL 536
Db 478 YIKFLKEVTNAWKDINKQFSRPTVEPMFVLRLVNLTRVADTLYKEKDITYSTAKGLKXNM 537
QY 537 IKSLLVYPM51 547
Db 538 INPILESVKI 548

RESULT 4

US-09-887-586A-52
; Sequence 52, Application US/09887586A
; Patent No. 6495354
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 64953541, Joseph P.
; APPLICANT: Starke, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/887,586A
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 52
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Lycopersicon esculentum
US-09-887-586A-52

Query Match 38.6%; Score 1123; DB 4; Length 548;

Best Local Similarity 38.8%; Pred. No. 1.9e-101;
Matches 214; Conservative 123; Mismatches 206; Indels 8; Gaps 3;

QY 2 MALTEEPKPIPIANFPPIISWGDQFLIYQKQVEQVQVNDLKKVEVRLKKEALDIPMKH 61
Db 1 MAASSADKCRPLANFHSVWGHFLSYTHETNQEKVEVDEYKETIRKMLVETCD---NS 57
QY 62 ANLLKLIDEIQRIGIPYHFEREIDHALQCIYETYGDNWNGDRS-----SLWFLMRKQGY 117
Db 58 TQKLVLDAMQRLGVAYHFDNEIETSIQNFIDASSKQNDNDNLYVWSLRFVLVQOQHY 117
QY 118 VTCDVFNKYKNGAPKQSLANDVEGLELYEATSMRVPGEIILEDALGFTSRSLIMTK 177
Db 118 MSSDVFKQFTNQDGKFKETLTNDVQGLLSLYEASHLVRNBEIEALTFTTTHLESIVS 177
QY 178 DAFSTNPALFTEIQRALKQPLMKRLPRIEAAQYIPYQOQDSHNKTLKLAKLEFNLLQS 237
Db 178 NLSNNNSLKVEGEALTQPTMTLPRMGARKYISYIENNDAAHLLHLKFKALDFNMLOK 237
QY 238 LHKEELSHVCXKWKAFDIKKNAPCLDRIVECYFWGLSGYEPQYSRARVFTKAVAVIT 297
Db 238 FHQRELSDLTRWKLDPANKYPYARDRLVECYFWILGVYFEPKYSRARKWMTKVLNLT 297
QY 298 LIDDTYDAYGTVEELKIITEAVERWSITCLDTPYMKPIYKLFMDTYTEMEEFKAGR 357
Db 298 IIDTDFAYATFDELVTFNDAIQRDWANAIDSQPMRPAYQALLDIYSEMEQVLSKCK 357
QY 358 TDLFNCCKEFVKEFVRLNLMVAKANE--GHPTTEHDPVVIITGGANLTTTCVLGMSD 416
Db 358 LDRVYAKNEMKVLVAYFKETQMLNDCDHPKYEQVENAIVSAGYMMISTTCLVGLIE 417
QY 417 IFTKESVEWAVSAPPLFRYSIGILGRRLNDLMTKAEQERKSSSSLESYMEKYNVEEYA 476
Db 418 FISHETFEWMNSESIVRASALIAMNDIVGHEDEQERGHVASLIECYMKDYGASKQET 477

QY 477 QTLIYKEVDVWKDINREYLTTKNIPRPLLMVAVIYLCQFLEVQYAGKDNFTRMGDEYKHL 536
Db 478 YIKFLKEVTNAWKDINKQFSRPTVEPMFVLRLVNLTRVADTLYKEKDITYSTAKGLKXNM 537
QY 537 IKSLLVYPM51 547
Db 538 INPILESVKI 548

RESULT 5

US-09-895-752-52
; Sequence 52, Application US/09895752
; Patent No. 6559297
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 65592971, Joseph P.
; APPLICANT: Starke, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/895,752
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 52
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Lycopersicon esculentum
US-09-895-752-52

Query Match 38.8%; Score 1123; DB 4; Length 548;

Best Local Similarity 38.8%; Pred. No. 1.9e-101;
Matches 214; Conservative 123; Mismatches 206; Indels 8; Gaps 3;

QY 2 MALTEEPKPIPIANFPPIISWGDQFLIYQKQVEQVQVNDLKKVEVRLKKEALDIPMKH 61
Db 1 MAASSADKCRPLANFHSVWGHFLSYTHETNQEKVEVDEYKETIRKMLVETCD---NS 57
QY 62 ANLLKLIDEIQRIGIPYHFEREIDHALQCIYETYGDNWNGDRS-----SLWFLMRKQGY 117
Db 58 TQKLVLDAMQRLGVAYHFDNEIETSIQNFIDASSKQNDNDNLYVWSLRFVLVQOQHY 117
QY 118 VTCDVFNKYKNGAPKQSLANDVEGLELYEATSMRVPGEIILEDALGFTSRSLIMTK 177
Db 118 MSSDVFKQFTNQDGKFKETLTNDVQGLLSLYEASHLVRNBEIEALTFTTTHLESIVS 177
QY 178 DAFSTNPALFTEIQRALKQPLMKRLPRIEAAQYIPYQOQDSHNKTLKLAKLEFNLLQS 237
Db 178 NLSNNNSLKVEGEALTQPTMTLPRMGARKYISYIENNDAAHLLHLKFKALDFNMLOK 237
QY 238 LHKEELSHVCXKWKAFDIKKNAPCLDRIVECYFWGLSGYEPQYSRARVFTKAVAVIT 297
Db 238 FHQRELSDLTRWKLDPANKYPYARDRLVECYFWILGVYFEPKYSRARKWMTKVLNLT 297
QY 298 LIDDTYDAYGTVEELKIITEAVERWSITCLDTPYMKPIYKLFMDTYTEMEEFKAGR 357
Db 298 IIDTDFAYATFDELVTFNDAIQRDWANAIDSQPMRPAYQALLDIYSEMEQVLSKCK 357
QY 358 TDLFNCCKEFVKEFVRLNLMVAKANE--GHPTTEHDPVVIITGGANLTTTCVLGMSD 416
Db 358 LDRVYAKNEMKVLVAYFKETQMLNDCDHPKYEQVENAIVSAGYMMISTTCLVGLIE 417
QY 417 IFTKESVEWAVSAPPLFRYSIGILGRRLNDLMTKAEQERKSSSSLESYMEKYNVEEYA 476
Db 418 FISHETFEWMNSESIVRASALIAMNDIVGHEDEQERGHVASLIECYMKDYGASKQET 477

QY 477 QTLIYKEVEDVWKDINREYLTNNIPRLLMAVILYLCQFLEVQYAGKDNFTRMGDEYKHL 536
Db 478 YIKFLKEVTNAWKDINKQFSRTEVPMFVLRLNTRVADTLYKEKDTYSTAKGKLKNN 537
QY 537 IKSLLYVPMI 547
Db 538 INPILIESVKI 548

RESULT 6
US-09-903-012B-52
; Sequence 52, Application US/09903012B
; Patent No. 6569656
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 6569656, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/903,012B
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 52
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Lycopersicon esculentum
US-09-903-012B-52

Query Match 38.6%; Score 1123; DB 4; Length 548;
Best Local Similarity 38.8%; Pred. No. 1.9e-101;
Matches 214; Conservative 123; Mismatches 206; Indels 8; Gaps 3;

QY 2 MALTEEKPRIPRIANPPISWGDFLIYQKQVEGVQIVNDLKEVRQLLKEALDIPMKH 61
Db 1 MAASSADKCRPLANFHPSPVWGVLHFLSYTHETNQEKVEVDEYKETIRKMLVETCD---NS 57
QY 62 ANLLKLIDEIQRIGIPYHFEREIDHALQCIYETYGDNWNGDRS----SLWFLMRKQGY 117
Db 58 TQKLVLDAMQRLGVAYHFDNEIETSIQNFIDASSKQNDNNDNLYVSLRFLVRQGHY 117
QY 118 VTCDVFNKYDXNGAFKQSLANDVEGLLELYEATSMRVPGGEIILEDALGFTSRSLSMTK 177
Db 118 MSSDVFKQFTNQDGKFKETLTNDVQGLLSLYEASHLRVRNEEILEEALTTTTHLESIVS 177
QY 178 DAFSTNPALFTEIQRALKQPLWKRLPRIEAAQYIPFYQQODSHNKTLLKLAKLEFNLLQS 237
Db 178 NLSNNNSLKVEVGEALTQPIRMTLPRMGARKYISYENNDAAHLLKFAKLDFFNLQK 237
QY 238 LHKEELSHVCKWKAFFDKKNAPCLDRIVECYFWGLSGYEPQYSRARVFTTKAVAVIT 297
Db 238 FHQRELSDLTRWKKLDLFANKYPYARDLVECYFWILGVYFEPKYSRARKMTKVLNLS 297
QY 298 LIDDTYDAGTYBELKIFTEAVERWSITCLDTPYMKPIYKLFMDYTEMEEFLAKEGR 357
Db 298 IIDDFTDAYATFDELVTNDAIQRWDANAIDSIOQYMRPAYQALLDIYSEMEQVLSKEGK 357
QY 358 TDLFNGCKEFVKEFVRNLVVEAKWANE-GHPTTEHDPPVITGGANLLTTTCYLGMSD 416
Db 358 LDRVYAKNEMKKLVRAFYKETQWLNDCDHI PKYEEQVENAIVSAGYMMISTTCLVGIE 417
QY 417 IFTKESVNAVASAPPLFRYSGILGRRLNDLMTKAEQERKHSSSLSLESYKENVNBEYA 476
Db 418 FISHEFEWLMNESVIVRASALIRAMNDIVGHEDEQERGHVASLIECYMKDYASKOET 477

QY 477 QTLIYKEVEDVWKDINREYLTNNIPRLLMAVILYLCQFLEVQYAGKDNFTRMGDEYKHL 536
Db 478 YIKFLKEVTNAWKDINKQFSRTEVPMFVLRLNTRVADTLYKEKDTYSTAKGKLKNN 537
QY 537 IKSLLYVPMI 547
Db 538 INPILIESVKI 548

RESULT 7
US-09-900-797-52
; Sequence 52, Application US/09900797
; Patent No. 6645762
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 6645762, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/900,797
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 52
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Lycopersicon esculentum
US-09-900-797-52

Query Match 38.6%; Score 1123; DB 4; Length 548;
Best Local Similarity 38.8%; Pred. No. 1.9e-101;
Matches 214; Conservative 123; Mismatches 206; Indels 8; Gaps 3;

QY 2 MALTEEKPRIPRIANPPISWGDFLIYQKQVEGVQIVNDLKEVRQLLKEALDIPMKH 61
Db 1 MAASSADKCRPLANFHPSPVWGVLHFLSYTHETNQEKVEVDEYKETIRKMLVETCD---NS 57
QY 62 ANLLKLIDEIQRIGIPYHFEREIDHALQCIYETYGDNWNGDRS----SLWFLMRKQGY 117
Db 58 TQKLVLDAMQRLGVAYHFDNEIETSIQNFIDASSKQNDNNDNLYVSLRFLVRQGHY 117
QY 118 VTCDVFNKYDXNGAFKQSLANDVEGLLELYEATSMRVPGGEIILEDALGFTSRSLSMTK 177
Db 118 MSSDVFKQFTNQDGKFKETLTNDVQGLLSLYEASHLRVRNEEILEEALTTTTHLESIVS 177
QY 178 DAFSTNPALFTEIQRALKQPLWKRLPRIEAAQYIPFYQQODSHNKTLLKLAKLEFNLLQS 237
Db 178 NLSNNNSLKVEVGEALTQPIRMTLPRMGARKYISYENNDAAHLLKFAKLDFFNLQK 237
QY 238 LHKEELSHVCKWKAFFDKKNAPCLDRIVECYFWGLSGYEPQYSRARVFTTKAVAVIT 297
Db 238 FHQRELSDLTRWKKLDLFANKYPYARDLVECYFWILGVYFEPKYSRARKMTKVLNLS 297
QY 298 LIDDTYDAGTYBELKIFTEAVERWSITCLDTPYMKPIYKLFMDYTEMEEFLAKEGR 357
Db 298 IIDDFTDAYATFDELVTNDAIQRWDANAIDSIOQYMRPAYQALLDIYSEMEQVLSKEGK 357
QY 358 TDLFNGCKEFVKEFVRNLVVEAKWANE-GHPTTEHDPPVITGGANLLTTTCYLGMSD 416
Db 358 LDRVYAKNEMKKLVRAFYKETQWLNDCDHI PKYEEQVENAIVSAGYMMISTTCLVGIE 417
QY 417 IFTKESVNAVASAPPLFRYSGILGRRLNDLMTKAEQERKHSSSLSLESYKENVNBEYA 476
Db 418 FISHEFEWLMNESVIVRASALIRAMNDIVGHEDEQERGHVASLIECYMKDYASKOET 477
QY 477 QTLIYKEVEDVWKDINREYLTNNIPRLLMAVILYLCQFLEVQYAGKDNFTRMGDEYKHL 536

Db 478 YIKFLKEVTNAWKDINKQFSRPTFVPMFVLRRVILNLTTRVADTLTYKEKDTYSTAKGKLKQM 537
QY 537 IKSLVYPMSI 547
Db 538 INPILIESVKI 548

RESULT 8

US-09-398-395A-32
; Sequence 32, Application US/09398395A
; Patent No. 6468772
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 64687721, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/398,395A
; CURRENT FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 556
; TYPE: PRT
; ORGANISM: Solanum tuberosum
US-09-398-395A-32

Query Match 35.8%; Score 1041; DB 4; Length 556;
Best Local Similarity 39.6%; Pred. No. 2.2e-93;
Matches 219; Conservative 105; Mismatches 209; Indels 20; Gaps 7;
QY 6 EEKPIRPIANFPSPISWGQDFLIY--OKVEQGVGQIVNDLKKEVRQLLKEALDIPMKHAN 63
Db 13 EEEIVRPIADFSPLWGRFHSFDNQIAGKYAQEIETLKEQSRILLSASSRRTL--AE 70
QY 64 LLKLIDEIQRIGIPVHFERIDHALQCIYET-----YGNWNGDRSSLSWFLMRKQG 115
Db 71 KLDLIDIVERLGIAHFHEKQIDDDLDQYKADPNFEAHEYNDL---QTLVQFRLLRQHG 127
QY 116 YYVTCDFVNNYKDNKGNAPKQSLANDVEGLLELYEATSMRVGPEIILEDALGTRSLSIM 175
Db 128 YNISPKLFIHQDAKGKFKESLNCNDIKGLNLNLYEASHVTRTHGEDILEEALAFSTAHL-- 185
QY 176 TKDAFSTNPALFTETORALKQPLWKELPRIEAAQVIPPYQQQDSHNKTLKLAKLEFNLL 235
Db 186 -SAAPHLKSPLSQVTHALEQSLHKSIPRVETRYFISYEEEBEQKNDVLLQFAKLDNFLL 244
QY 236 QSLHKEELSHVCKWKWKAFDIKKNAPCLDRIVECYFWGLSGVGEPOYSRARVFTKAVAV 295
Db 245 QMLHKQELSEVSRWKKDLDFVTLLPYARDRAVECFWTVGVAEPQYSQARVWLAKTIAM 304
QY 296 ITLDDTYDAYGTVEELKIFTEAVERWSITCLDTPLEYMKPIYKLFMDTYTEMEEFLEKE 355
Db 305 ISIVDDTFDAYGIVKLEIYTDIAQRWDISQIDRLPDYMKISYKALLDLYNDYEMELSKD 364
QY 356 GRDLFCNGCKEFVKEFVRNLNLYEAKWANEHGHPTTEHDPVVIITGGANLLTTTCVLGMS 415
Db 365 GRSDVHVYAKERKMEIVRNYPVEAKWFTIEGYPPVSEYLSNALATSTYLLTTTSYLGMK 424
QY 416 DIFTKESVEWAVASAPPLPRYSGLGRNLNDLTHKAEQERKSSSLESYMKYNNVEEY 475
Db 425 SA-NKQDFEWLAKNPKILEANVTLCRVDDIATYVEKGRGQIATGIECYMRDYGVSTEK 493
QY 476 AQTLYIKEYEDVWKDINREYLTATKNIPRELLMAVLYLCQFLEVQYA-GKDNFTRMGDYK 534

Db 484 AMEKFOEMAETAWKQVNEGILRTPVSTETILTRILNLARIIDVTYKHNQDGYTHPEKVLK 543
QY 535 HLIKSLLVYPMSI 547
Db 544 PHIALLVDSIEI 556

RESULT 9

US-09-887-586A-32
; Sequence 32, Application US/09887586A
; Patent No. 6496354
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 64963541, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/887,586A
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 556
; TYPE: PRT
; ORGANISM: Solanum tuberosum
US-09-887-586A-32

Query Match 35.8%; Score 1041; DB 4; Length 556;
Best Local Similarity 39.6%; Pred. No. 2.2e-93;
Matches 219; Conservative 105; Mismatches 209; Indels 20; Gaps 7;
QY 6 EEKPIRPIANFPSPISWGQDFLIY--OKVEQGVGQIVNDLKKEVRQLLKEALDIPMKHAN 63
Db 13 EEEIVRPIADFSPLWGRFHSFDNQIAGKYAQEIETLKEQSRILLSASSRRTL--AE 70
QY 64 LLKLIDEIQRIGIPVHFERIDHALQCIYET-----YGNWNGDRSSLSWFLMRKQG 115
Db 71 KLDLIDIVERLGIAHFHEKQIDDDLDQYKADPNFEAHEYNDL---QTLVQFRLLRQHG 127
QY 116 YYVTCDFVNNYKDNKGNAPKQSLANDVEGLLELYEATSMRVGPEIILEDALGTRSLSIM 175
Db 128 YNISPKLFIHQDAKGKFKESLNCNDIKGLNLNLYEASHVTRTHGEDILEEALAFSTAHL-- 185
QY 176 TKDAFSTNPALFTETORALKQPLWKELPRIEAAQVIPPYQQQDSHNKTLKLAKLEFNLL 235
Db 186 -SAAPHLKSPLSQVTHALEQSLHKSIPRVETRYFISYEEEBEQKNDVLLQFAKLDNFLL 244
QY 236 QSLHKEELSHVCKWKWKAFDIKKNAPCLDRIVECYFWGLSGVGEPOYSRARVFTKAVAV 295
Db 245 QMLHKQELSEVSRWKKDLDFVTLLPYARDRAVECFWTVGVAEPQYSQARVWLAKTIAM 304
QY 296 ITLDDTYDAYGTVEELKIFTEAVERWSITCLDTPLEYMKPIYKLFMDTYTEMEEFLEKE 355
Db 305 ISIVDDTFDAYGIVKLEIYTDIAQRWDISQIDRLPDYMKISYKALLDLYNDYEMELSKD 364
QY 356 GRDLFCNGCKEFVKEFVRNLNLYEAKWANEHGHPTTEHDPVVIITGGANLLTTTCVLGMS 415
Db 365 GRSDVHVYAKERKMEIVRNYPVEAKWFTIEGYPPVSEYLSNALATSTYLLTTTSYLGMK 424
QY 416 DIFTKESVEWAVASAPPLPRYSGLGRNLNDLTHKAEQERKSSSLESYMKYNNVEEY 475
Db 425 SA-NKQDFEWLAKNPKILEANVTLCRVDDIATYVEKGRGQIATGIECYMRDYGVSTEK 483
QY 476 AQTLYIKEYEDVWKDINREYLTATKNIPRELLMAVLYLCQFLEVQYA-GKDNFTRMGDYK 534
Db 484 AMEKFOEMAETAWKQVNEGILRTPVSTETILTRILNLARIIDVTYKHNQDGYTHPEKVLK 543

```
QY 535 HLIKSLLVPMISI 547
Db 544 PHIALLVDSIEI 556

RESULT 10
US-09-895-752-32
; Sequence 32, Application US/09895752
; Patent No. 6559297
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 6559297, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/895,752
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 556
; TYPE: PRT
; ORGANISM: Solanum tuberosum
US-09-895-752-32

Query Match 35.8%; Score 1041; DB 4; Length 556;
Best Local Similarity 39.6%; Pred. No. 2.2e-93;
Matches 219; Conservative 105; Mismatches 209; Indels 20; Gaps 7;

QY 6 EEKPIRPIANFPSPISWGDOFLY--OKQVEQVQIYNLKKVRLQKLEALDIPMKHAN 63
Db 13 EEEIVRPIADPSPISWGDRFHSFLDNQIAGKYAQEIETLKEQSRILSASSRRL--AE 70
QY 64 LKLIDEIQRIGIPHFEREIDHALQCIYET-----YGDNMNGDRSLWFLMRKQG 115
Db 71 KLDLIDIVERLGIAYHFEXKQIDDMLDQFYKADPNFEAHEYNDL---QTLVSVQFRLLRQH 127
QY 116 YVTCDFVNNYKQNGAFKQSLANDVEGLLEYEATSMRVPGIEIILEDALGFTSRSLIM 175
Db 128 YNISPKLFIRFQDAKGFESLNCNDIKGLNLNLYEASHVRTHGEDILEEALAFSTAHL 185
QY 176 TKDAFSTNPALFTEIQRALKQPLMKRLPRIEAAQYIPFYQQQDSHNKTLKLAKLEFNLL 235
Db 186 -SAAPHLKSPLSQVTHALEQSLHKSIPRVETRYFISYEEBEEQKNDVLLQFALDFNLL 244
QY 236 QSLHKEELSHVCKWKAQFIDKQAPCLDRIVECYFWGLSGYEPQYSRARVFTKAVAV 295
Db 245 QMLHKEELSEVRWKKLDFTVTLPYARDRAVECYFTWGVYAEQYSQARVMLAKTIAM 304
QY 296 ITLIDDTYDAYGTVEELKIFTEAVERWSITCLDTLPEYMKPIYKLFMDTYTEMEEF 355
Db 305 ISIVDDTFDAYGIVKELEIYTDIQWDSIQIDRLPDYMKISYKALLDLYNDYEMELSKD 364
QY 356 GRDLPNCGKEFVKFVRNLMVEAKWANEGHPTTEEDHPVVIITGGANLLTTTCYLGMS 415
Db 365 GRSDVVHYAKERMEIVRNRYFVEAKFIEGMPVSEYLSNALATSYLLTTTSYLGMK 424
QY 416 DIFTKESVWAVSAPPLFRYSGLILRRNLDMTHKAEQERKSSSSLESYMKENYNEEY 475
Db 425 SA-NKQDFEMLAKNPXILEANVTLCRVDDIATYVEKGRGOIATGIECYMRDYGVST 483
QY 476 AQTLIYKEVEDWVDKINREYLTNKPRLMAVILYLCQFLEVQYA-GKDNFTRMGDEYK 534
Db 483
```

```
Db 484 AMEKFOEMAETAWKDVNEGILRPTPVSTELTRILNARIIDVTYKHNQDGYTHPEKVLK 543
QY 535 HLIKSLLVPMISI 547
Db 544 PHIALLVDSIEI 556

RESULT 11
US-09-903-012B-32
; Sequence 32, Application US/09903012B
; Patent No. 6569656
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 6569656, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/903,012B
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 556
; TYPE: PRT
; ORGANISM: Solanum tuberosum
US-09-903-012B-32

Query Match 35.8%; Score 1041; DB 4; Length 556;
Best Local Similarity 39.6%; Pred. No. 2.2e-93;
Matches 219; Conservative 105; Mismatches 209; Indels 20; Gaps 7;

QY 6 EEKPIRPIANFPSPISWGDOFLY--OKQVEQVQIYNLKKVRLQKLEALDIPMKHAN 63
Db 13 EEEIVRPIADPSPISWGDRFHSFLDNQIAGKYAQEIETLKEQSRILSASSRRL--AE 70
QY 64 LKLIDEIQRIGIPHFEREIDHALQCIYET-----YGDNMNGDRSLWFLMRKQG 115
Db 71 KLDLIDIVERLGIAYHFEXKQIDDMLDQFYKADPNFEAHEYNDL---QTLVSVQFRLLRQH 127
QY 116 YVTCDFVNNYKQNGAFKQSLANDVEGLLEYEATSMRVPGIEIILEDALGFTSRSLIM 175
Db 128 YNISPKLFIRFQDAKGFESLNCNDIKGLNLNLYEASHVRTHGEDILEEALAFSTAHL 185
QY 176 TKDAFSTNPALFTEIQRALKQPLMKRLPRIEAAQYIPFYQQQDSHNKTLKLAKLEFNLL 235
Db 186 -SAAPHLKSPLSQVTHALEQSLHKSIPRVETRYFISYEEBEEQKNDVLLQFALDFNLL 244
QY 236 QSLHKEELSHVCKWKAQFIDKQAPCLDRIVECYFWGLSGYEPQYSRARVFTKAVAV 295
Db 245 QMLHKEELSEVRWKKLDFTVTLPYARDRAVECYFTWGVYAEQYSQARVMLAKTIAM 304
QY 296 ITLIDDTYDAYGTVEELKIFTEAVERWSITCLDTLPEYMKPIYKLFMDTYTEMEEF 355
Db 305 ISIVDDTFDAYGIVKELEIYTDIQWDSIQIDRLPDYMKISYKALLDLYNDYEMELSKD 364
QY 356 GRDLPNCGKEFVKFVRNLMVEAKWANEGHPTTEEDHPVVIITGGANLLTTTCYLGMS 415
Db 365 GRSDVVHYAKERMEIVRNRYFVEAKFIEGMPVSEYLSNALATSYLLTTTSYLGMK 424
QY 416 DIFTKESVWAVSAPPLFRYSGLILRRNLDMTHKAEQERKSSSSLESYMKENYNEEY 475
Db 425 SA-NKQDFEMLAKNPXILEANVTLCRVDDIATYVEKGRGOIATGIECYMRDYGVST 483
QY 476 AQTLIYKEVEDWVDKINREYLTNKPRLMAVILYLCQFLEVQYA-GKDNFTRMGDEYK 534
Db 484 AMEKFOEMAETAWKDVNEGILRPTPVSTELTRILNARIIDVTYKHNQDGYTHPEKVLK 543
```

QY 535 HLKSLVVPMSI 547
Db 544 PHIALVDSIEI 556

RESULT 12

US-09-900-797-32
; Sequence 32, Application US/09900797
; Patent No. 6645762
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 66457621, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/900,797
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 556
; TYPE: PRT
; ORGANISM: Solanum tuberosum
US-09-900-797-32

Query Match 35.8%; Score 1041; DB 4; Length 556;
Best Local Similarity 39.6%; Pred. No. 2.2e-93;
Matches 219; Conservative 105; Mismatches 209; Indels 20; Gaps 7;

QY 6 BEKTRPIANFPSSWGDQFLIY--QKQVEQVEQIVNDLKKEVRQLLKEALDIPMKHAN 63
Db 13 EEEIVRPVADSPSLWGDQFLSPSIKNQVAEYAEIEALKQETRNWL---LATGMKLAD 66
QY 64 LKLDIDEIQRIGIPYHFEREIDHALQCIYET-----YGDWNGDRSSLMFLMRKQG 115
Db 71 KLDLDIVERLGIAYHFEQIDMDLQFYKADPNFEAHEYNDL---QTLVQFLLRQHG 127
QY 116 YVYVCDVFNKYKNGAFKQSLANDVEGLELEYEATSMRVPGIIEEDALGFTSRSLSIM 175
Db 128 YNISPKLFIREFODAKGFKESLNDIKGLNLNLYEASHVRTHGEDILBEALAFSTAHL-- 185
QY 176 TKDAPSTNPALEIQRALKQPLWKRLPRIEAAQYIPFYQQQDSHNKTLKLAKLEFN 235
Db 186 -SAAPHLKSPKQVTHALEQSLHSIPRVETRYFISYIEEEQKNDVLLQFALDFNLL 244
QY 236 QSLHKEELSHVCKWKAFDIKKNAPCLDRDRIVECYFWGLSGYEPQYSRARVFTKAVAV 295
Db 245 QMLHKEELSEVSWKDLDFVTLTPYARDRAVECYFWTMGVYAEQYSQARVMLAKTIAM 304
QY 296 ITLIDTIDYDAYGYEELKFTTEAVERNSITCLDTPLEPMKPIYKLFWDITYTEMEEF 355
Db 305 ISIVDDTIDYDAYGIVKELEYTDIAQRWDISQIDRLPDYMKISYKALLDLYNDEMELSKD 364
QY 356 GRITDLENGCKEFKVFERNLMVEAKWANEHGIPTTEBHDPPVITGGANLLTTTCYLGMS 415
Db 365 GRSDVVHYAKRMKEIVRNYPVFAKMFIEGYMPPVSEYLSNALATSTYLLTTTTSYLG 424
QY 416 DIFTKESVEMWASAPPLFRYSIGLGRNLNLMTHKAEQERKHSSESYMKEYNNVEEY 475
Db 425 SA-NKQDFWLAKNPKILEANVILCRVIDDIAIYEVKEKGQIATGECWYRDYGVSTEK 483
QY 476 AQTLLIYKEVEDWVKDINREYLTNNIPRPLMAVILCOFLEVQYA--GKDNFTRMGDEYK 534
Db 484 AMEKFOEMAETAAMKVDYNEGILRTPPVSTETILTRILNLAIRIDVTYKKNQDGYTHPEKVK 543

QY 535 HLKSLVVPMSI 547
Db 544 PHIALVDSIEI 556

RESULT 13

US-09-398-395A-2
; Sequence 2, Application US/09398395A
; Patent No. 6468772
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 64687721, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/398,395A
; CURRENT FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Nicotiana tabacum
US-09-398-395A-2

Query Match 35.1%; Score 1020.5; DB 4; Length 548;
Best Local Similarity 40.0%; Pred. No. 2.3e-91;
Matches 222; Conservative 102; Mismatches 202; Indels 29; Gaps 9;

QY 6 BEKTRPIANFPSSWGDQFLIY--QKQVEQVEQIVNDLKKEVRQLLKEALDIPMKHAN 63
Db 10 EEEIVRPVADSPSLWGDQFLSPSIKNQVAEYAEIEALKQETRNWL---LATGMKLAD 66
QY 64 LKLDIDEIQRIGIPYHFEREIDHALQCIYETVYGDWNGD-----RSSLMFLMRKQGYV 118
Db 67 TLNLDITIERLGISYHFEKEDIDLDQIY-----NQNSNCNDLCTSLAQFRLLRQHGFI 121
QY 119 TCDVFNKYKNGAFKQSLANDVEGLELEYEATSMRVPGIIEEDALGFTSRSLSIMTKD 178
Db 122 SPEIFSKQDENGKFKESLADSVLGLNLNLYEASHVRTHADDILEDALAFSTIHLE----- 176
QY 179 AFTSNPALFT---EIQRALKQPLWKRLPRIEAAQVI-PFYQQQDSHNKTLKLAKLEFN 233
Db 177 -SAAPHLKSPLEQVTHALQCLHKGVPVTRFPISSYDKQSKNNVLLRFAKLDFN 234
QY 234 LLOSLHKEELSHVCKWKAFDIKKNAPCLDRDRIVECYFWGLSGYEPQYSRARVFTKAV 293
Db 235 LLOMLHKEELAQVSRWKDLDFVTLTPYARDRAVECYFWALGVYFEPQYSQARVMLVKT 294
QY 294 AVITLIDTIDYDAYGYEELKFTTEAVERNSITCLDTPLEPMKPIYKLFWDITYTEMEEF 353
Db 295 SMISIVDDTIDYDAYGIVKELEYTDIAQRWDINEIDRLPDYMKISYKAILDLYKDYEKELS 354
QY 354 KEGRTDLENGCKEFKVFERNLMVEAKWANEHGIPTTEBHDPPVITGGANLLTTTCYLG 413
Db 355 SAGSHIYCHAIERKKEVVRNRYNVVESTFIEGYTPVSEYLSNALATTTTYLLATTSYLG 414
QY 414 MSDIFTKESVEMWASAPPLFRYSIGLGRNLNLMTHKAEQERKHSSESYMKEYNNVE 473
Db 415 MKSA-TEQDFEWLSKNPKILEASVILCRVIDDIAIYEVKEKRGQIATGIECCNRDYGIST 473
QY 474 EYAQTLIYKEVEDWVKDINREYLTNNIPRPLMAVILCOFLEVQYA--GKDNFTRMGDE 532
Db 474 KEAMAKFOEMAETAAMKVDYNEGILRTPPVSTETILTRILNLAIRIVEYTHNLDGYTHPEKV 533
QY 533 YKHLKSLVVPMSI 547

```
Db 534 LKPHIINLVDSIKI 548
RESULT 14
US-09-887-586A-2
; Sequence 2, Application US/09887586A
; Patent No. 6495354
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 6495354, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/887,586A
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Nicotiana tabacum
US-09-887-586A-2

Query Match 35.1%; Score 1020.5; DB 4; Length 548;
Best Local Similarity 40.0%; Pred. No. 2.3e-91;
Matches 222; Conservative 102; Mismatches 202; Indels 29; Gaps 9;

Qy 6 BEKPIRPIANPPSINGDQFLY--QKQVEQGVQIIVNDLKEVQQLLKEALDIPMKHAN 63
Db 10 EEEIVRPVADPSLWGDQFLSFKNOVAEYQAEIALKEQTRNML---LATGMKLAD 66
Qy 64 LKLDLDELQIGIYHFEREDHALOCIVETYGDNWGD-----RSSLWFLMRKOGYVV 118
Db 67 TLNLIDTIERLGISYHFEKIDIDLOIY-----NQNNSCNDLCTSAQFLRLRQHGFI 121
Qy 119 TCDVFNKYKQNGAFKQSLANDVEGLELYEATSMRVPGIEIILEDALGFRSRLSINTKD 178
Db 122 SPEIFSKFODENGKFKESLASDVLGLNLYEASHVTHADDDILEDALAFSTHLE----- 176
Qy 179 AFTSNPALFT---EIQRAKQPLWKRLPRIEAAQYI-PFYQQODSHNKTLLKLAKLEFN 233
Db 177 --SAAPHLKSPLEQVTHALEQCLHKGVPVETRFIFSIYDKESQSKNNVLLRFAKLDFN 234
Qy 234 LLOSLHKEELSHVCKWKAFFDIKKNAPCLDRIVECYFWGLSGYEPQYSRABVFTKAV 293
Db 235 LLOMLHKQELAQVSRWKKDLDFVTTLPYARDRVVECYFWALGVYFEPQYSQARVMLVKT 294
Qy 294 AVITLIDDTYDAYGYEELKIIFTEAVERWSITCLDTLPYMKPIYKLFMDTYTEMEEF 353
Db 295 SMSIVDDTFDAYGTVEKLEAYTDAIQRDWINEIDRLPDYMKISYKAILDLYKDYKELS 354
Qy 354 KEGRTDLFNCGEFVKFVVRNLMVEAKWANEHIPPTEEDHPVVIITGGANLTTTCYLG 413
Db 355 SAGRSHIVCHATERMKEVRNVNVESTWTFIEGYTPPVSEYLSNALATTTTYYLATTSYLG 414
Qy 414 MSDIFTKESVEMAVSAPPLFRYSGILGRRLNDLMTKAEQERKSSSSLESYKENVNNE 473
Db 415 MKSA-TEQDFSLKSNPKILEASVIIICRVDDTATVEKSRGQIATGECNRDYGIST 473
Qy 474 EYAQTLIYKEVEDVWKDINREYLTNNIPRPLMAVILYLCQFLEVOYA-GKDNFTRMGE 532
Db 474 KZAMAKFQNAETAWKDINEGLRPTVPVSTFETPLNLARIVEVYIHNLDGTYHPEKV 533
Qy 533 YKHLIKSLLYVPMIS 547

Query Match 35.1%; Score 1020.5; DB 4; Length 548;
Best Local Similarity 40.0%; Pred. No. 2.3e-91;
Matches 222; Conservative 102; Mismatches 202; Indels 29; Gaps 9;

Qy 6 BEKPIRPIANPPSINGDQFLY--QKQVEQGVQIIVNDLKEVQQLLKEALDIPMKHAN 63
Db 10 EEEIVRPVADPSLWGDQFLSFKNOVAEYQAEIALKEQTRNML---LATGMKLAD 66
Qy 64 LKLDLDELQIGIYHFEREDHALOCIVETYGDNWGD-----RSSLWFLMRKOGYVV 118
Db 67 TLNLIDTIERLGISYHFEKIDIDLOIY-----NQNNSCNDLCTSAQFLRLRQHGFI 121
Qy 119 TCDVFNKYKQNGAFKQSLANDVEGLELYEATSMRVPGIEIILEDALGFRSRLSINTKD 178
Db 122 SPEIFSKFODENGKFKESLASDVLGLNLYEASHVTHADDDILEDALAFSTHLE----- 176
Qy 179 AFTSNPALFT---EIQRAKQPLWKRLPRIEAAQYI-PFYQQODSHNKTLLKLAKLEFN 233
Db 177 --SAAPHLKSPLEQVTHALEQCLHKGVPVETRFIFSIYDKESQSKNNVLLRFAKLDFN 234
Qy 234 LLOSLHKEELSHVCKWKAFFDIKKNAPCLDRIVECYFWGLSGYEPQYSRABVFTKAV 293
Db 235 LLOMLHKQELAQVSRWKKDLDFVTTLPYARDRVVECYFWALGVYFEPQYSQARVMLVKT 294
Qy 294 AVITLIDDTYDAYGYEELKIIFTEAVERWSITCLDTLPYMKPIYKLFMDTYTEMEEF 353
Db 295 SMSIVDDTFDAYGTVEKLEAYTDAIQRDWINEIDRLPDYMKISYKAILDLYKDYKELS 354
Qy 354 KEGRTDLFNCGEFVKFVVRNLMVEAKWANEHIPPTEEDHPVVIITGGANLTTTCYLG 413
Db 355 SAGRSHIVCHATERMKEVRNVNVESTWTFIEGYTPPVSEYLSNALATTTTYYLATTSYLG 414
Qy 414 MSDIFTKESVEMAVSAPPLFRYSGILGRRLNDLMTKAEQERKSSSSLESYKENVNNE 473
Db 415 MKSA-TEQDFSLKSNPKILEASVIIICRVDDTATVEKSRGQIATGECNRDYGIST 473
Qy 474 EYAQTLIYKEVEDVWKDINREYLTNNIPRPLMAVILYLCQFLEVOYA-GKDNFTRMGE 532
Db 474 KZAMAKFQNAETAWKDINEGLRPTVPVSTFETPLNLARIVEVYIHNLDGTYHPEKV 533
Qy 533 YKHLIKSLLYVPMIS 547
```

Db 534 LKPHIINLLVDSIKI 548

Search completed: June 3, 2004, 16:34:05
Job time : 31 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
OM protein - protein search, using sw model
Run on: June 3, 2004, 16:32:55 ; Search time 50 Seconds
(without alignments)
3089.099 Million cell updates/sec

Title: US-09-763-822A-14
Perfect score: 2906
Sequence: 1 TMLTEKPIRPIANFPSPS.....GDEYKHLKSLVYPMSIGS 549

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 1155919 seqs, 281338677 residues
Total number of hits satisfying chosen parameters: 1155919
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2890	99.4	548	15	US-10-411-066-38
2	1462.5	50.3	551	15	US-10-411-066-40
3	1123	38.6	548	9	US-09-887-586A-52
4	1123	38.6	548	9	US-09-903-012-52
5	1123	38.6	548	10	US-09-900-797-52
6	1123	38.6	548	12	US-09-893-820-52
7	1099.5	37.8	567	12	US-10-424-599-186274
8	1041	35.8	556	9	US-09-887-586A-32
9	1041	35.8	556	9	US-09-903-012-32
10	1041	35.8	556	10	US-09-900-797-32
11	1041	35.8	556	12	US-09-893-820-32
12	1020.5	35.1	548	9	US-09-887-586A-2
13	1020.5	35.1	548	9	US-09-903-012-2
14	1020.5	35.1	548	10	US-09-900-797-2
15	1020.5	35.1	548	12	US-09-893-820-2

16	1017.5	35.0	548	9	US-09-887-586A-8	Sequence 8, Appli
17	1017.5	35.0	548	9	US-09-903-012-8	Sequence 8, Appli
18	1017.5	35.0	548	10	US-09-900-797-8	Sequence 8, Appli
19	1017.5	35.0	548	12	US-09-893-820-8	Sequence 8, Appli
20	1010.5	34.8	548	9	US-09-887-586A-6	Sequence 6, Appli
21	1010.5	34.8	548	9	US-09-903-012-6	Sequence 6, Appli
22	1010.5	34.8	548	10	US-09-900-797-6	Sequence 6, Appli
23	1010.5	34.8	548	12	US-09-893-820-6	Sequence 6, Appli
24	1009.5	34.7	548	9	US-09-887-586A-12	Sequence 12, Appl
25	1009.5	34.7	548	9	US-09-903-012-12	Sequence 12, Appl
26	1009.5	34.7	548	10	US-09-900-797-12	Sequence 12, Appl
27	1009.5	34.7	548	12	US-09-893-820-12	Sequence 12, Appl
28	1001.5	34.5	548	9	US-09-887-586A-10	Sequence 10, Appl
29	1001.5	34.5	548	9	US-09-903-012-10	Sequence 10, Appl
30	1001.5	34.5	548	10	US-09-900-797-10	Sequence 10, Appl
31	1001.5	34.5	548	12	US-09-893-820-10	Sequence 10, Appl
32	1000.5	34.4	548	9	US-09-887-586A-4	Sequence 4, Appli
33	1000.5	34.4	548	9	US-09-903-012-4	Sequence 4, Appli
34	1000.5	34.4	548	10	US-09-900-797-4	Sequence 4, Appli
35	1000.5	34.4	548	12	US-09-893-820-4	Sequence 4, Appli
36	967.5	33.3	601	9	US-09-887-586A-42	Sequence 42, Appl
37	967.5	33.3	601	9	US-09-903-012-42	Sequence 42, Appl
38	967.5	33.3	601	10	US-09-900-797-42	Sequence 42, Appl
39	967.5	33.3	601	12	US-09-893-820-42	Sequence 42, Appl
40	888	30.6	558	12	US-10-425-114-49923	Sequence 49923, A
41	853	29.4	550	10	US-09-860-282A-7	Sequence 7, Appli
42	853	29.4	550	10	US-09-860-282A-17	Sequence 17, Appl
43	852	29.3	550	9	US-09-887-586A-28	Sequence 28, Appl
44	852	29.3	550	9	US-09-903-012-28	Sequence 28, Appl
45	852	29.3	550	10	US-09-900-797-28	Sequence 28, Appl

ALIGNMENTS

RESULT 1

US-10-411-066-38
; Sequence 38, Application US/10411066
; Publication No. US20040005678A1
; GENERAL INFORMATION:
; APPLICANT: KEASLING, JAY
; APPLICANT: MARTIN, VINCENT
; APPLICANT: PITERA, DOUGLAS
; APPLICANT: WITHERS III, SYDNOR T.
; APPLICANT: NEWMAN, JACK
; TITLE OF INVENTION: BIOSYNTHESIS OF AMORPHA-4,11-DIENE
; FILE REFERENCE: 2000-0007.20
; CURRENT APPLICATION NUMBER: US/10/411,066
; CURRENT FILING DATE: 2000-04-09
; PRIOR APPLICATION NUMBER: 10/006,909
; PRIOR FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 38
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Artemisia annua
US-10-411-066-38

Query Match	99.4%	Score 2890;	DB 15;	Length 548;
Best Local Similarity	99.6%	Pred No. 3.3e-259;		
Matches 546;	Conservative 1;	Mismatches 1;	Indels 0;	Gaps 0;
Qy	2	MALTEKPIRPIANFPPIGWGDFLIYKQVQGVQVEQVQVNDLKKEVRQLLKEALDIPMKH 61		
Db	1	MALTEKPIRPIANFPPIGWGDFLIYKQVQGVQVEQVQVNDLKKEVRQLLKEALDIPMKH 60		
Qy	62	ANLKLDEIORLQIPYHFEIDHALQCIYETGYDGNWNGDRSSLMFRLMRKQYVTC 121		
Db	61	ANLKLDEIORLQIPYHFEIDHALQCIYETGYDGNWNGDRSSLMFRLMRKQYVTC 120		
Qy	122	VFNVKNKGNGAFKQSLANDVEGLLELYEATSMRVFGEILLEDAIGFTRGLSMTKDAPS 181		

Db 121 VFNNYKDKNGAFKQSLANDVEGELLEZEATSMRVPGELIILEDALGFTSRSLINTKDAFS 180
Qy 182 TNPALFTEIQORALKQPIWKLPLRIEAAQYIPFYQQQDSHNKTLKLAKLEFNLLQSLHKE 241
Db 181 TNPALFTEIQORALKQPIWKLPLRIEAAQYIPFYQQQDSHNKTLKLAKLEFNLLQSLHKE 240
Qy 242 ELSHVCKWKAPDIKKNAPCLDRIVECYFWGLSGGYEPOYSRAREFTTKAVAVITLIDD 301
Db 241 ELSHVCKWKAPDIKKNAPCLDRIVECYFWGLSGGYEPOYSRAREFTTKAVAVITLIDD 300
Qy 302 TYDAYGTVEELKIPTFAVERWSITCLDTPYMKPIYKLFMDYVTEMEEFLLAKEGRDLE 361
Db 301 TYDAYGTVEELKIPTFAVERWSITCLDTPYMKPIYKLFMDYVTEMEEFLLAKEGRDLE 360
Qy 362 NCGKEPVKEFVRLNLMVEAKWANEHGIPTTEHDPVVIITGGANLLTTTCVYLGMSDIFTKE 421
Db 361 NCGKEPVKEFVRLNLMVEAKWANEHGIPTTEHDPVVIITGGANLLTTTCVYLGMSDIFTKE 420
Qy 422 SVEWAVSAPPLFRYSIGILGRRLNLMTHKAEQERKHSSSSLESYMKYNNVEEYAOQLIY 481
Db 421 SVEWAVSAPPLFRYSIGILGRRLNLMTHKAEQERKHSSSSLESYMKYNNVEEYAOQLIY 480
Qy 482 KEVEDVWKDINREYLTITKNIIPRLMAVIYLCQFLEVQYAGKDNFTRMGDEYKHLIKSL 541
Db 481 KEVEDVWKDINREYLTITKNIIPRLMAVIYLCQFLEVQYAGKDNFTRMGDEYKHLIKSL 540
Qy 542 VYPMSIGS 549
Db 541 VYPMSIGS 548

RESULT 2

US-10-411-066-40
; Sequence 40, Application US/10411066
; Publication No. US20040005678A1
; GENERAL INFORMATION:
; APPLICANT: KEASLING, JAY
; APPLICANT: MARTIN, VINCENT
; APPLICANT: PITERA, DOUGLAS
; APPLICANT: WITHERS III, SYDNOR T.
; APPLICANT: NEWMAN, JACK
; TITLE OF INVENTION: BIOSYNTHESIS OF AMORPHA-4,11-DIENE
; FILE REFERENCE: 2000-0007-20
; CURRENT APPLICATION NUMBER: US/10/411,066
; CURRENT FILING DATE: 2000-04-09
; PRIOR APPLICATION NUMBER: 10/006,909
; PRIOR FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 551
; TYPE: PRT
; ORGANISM: Artemisia annua
US-10-411-066-40

Query Match 50.3%; Score 1462.5; DB 15; Length 551;
Best Local Similarity 52.0%; Pred. No. 1.2e-126;
Matches 282; Conservative 97; Mismatches 162; Indels 1; Gaps 1;
Qy 2 MALTEEKPIRPIANPPSIWGOFLIYQKQVGEQVQVNDLKKEVRQLLKALDIPMKH 61
Db 1 MSLIVEDVRNANPSEIWGOFLAYDQDEGEVQVNDLKKEVKSLLTALNSPTQH 60
Qy 62 ANLLKLIDEIQRLGIPYHFEREIDHALQCIYETTYGDNWNGDRSSSLWFLMRKQGYVTC 121
Db 61 TELLKFIDAERLGIAFYEEBINGVQFMYTAYGDKMTGGNTSLWFLMRQHGPFVSSD 120
Qy 122 VFNNYKDKNGAFKQSLANDVEGELLEZEATSMRVPGELIILEDALGFTSRSLINTKDAFS 181
Db 121 IFSTYKDEGRFKEKLEKDVHGLLEZEAYAFVPGEGILDDALVFRTCLDEIAKNPSL 180
Qy 182 TNPALFTEIQORALKQPIWKLPLRIEAAQYIPFYQQQDSHNKTLKLAKLEFNLLQSLHKE 241

Db 181 SNSAVSSQIRBALQPLHKLRLPRLEALRYIPFYQQQASHSETLLKLAKLGFNQLQSLHKK 240
Qy 242 ELSHVCKWKAPDIKKNAPCLDRIVECYFWGLSGGYEPOYSRAREFTTKAVAVITLIDD 301
Db 241 ELSIISKWKSFDVANNLPYARNRPVECYFWALAVYFEPQYSESRVLSRFSIQITFLDD 300
Qy 302 TYDAYGTVEELKIPTFAVERWSITCLDTPYMKPIYKLFMDYVTEMEEFLLAKEGRDLE 361
Db 301 TYDAYGTVEELQEFTQATQWISITCLDGLPESMKLIFQMLVKIFEBIEBILSKDGKQHHV 360
Qy 362 NCGKEPVKEFVRLNLMVEAKWANEHGIPTTEHDPVVIITGGANLLTTTCVYLGMSDIFTKE 421
Db 361 NYIKETLKEAVQSYMTENARNAKEVYPIPIEHTKVSYSIGYKALAVAGACMGDVIADD 420
Qy 422 SVEWAVSAPPLFRYSIGILGRRLNLMTHKAEQERKHSSSSLESYMKYNNVEEYAOQLIY 481
Db 421 SFEWVFTNPPLVNAACLLCRTMDDLGLSHKGEQDRKHVASTIECYMKQPDASEQOAYESLN 480
Qy 482 KEVEDVWKDINREYLTITKNIIPRLMAVIYLCQFLEVQYAGKDNFTRMGDEYKHLIKSL 540
Db 481 KKVEDAWKEINREMITCKOYNIHVAMRVLFNRSRSDVLYKKNKDHFTHVGEVYINHIKSL 540
Qy 541 LV 542
Db 541 FV 542

RESULT 3

US-09-887-586A-52
; Sequence 52, Application US/09887586A
; Patent No. US20020094556A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20020094556A1, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/887,586A
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 52
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Lycopersicon esculentum
US-09-887-586A-52

Query Match 38.6%; Score 1123; DB 9; Length 548;
Best Local Similarity 38.8%; Pred. No. 4e-95;
Matches 214; Conservative 123; Mismatches 206; Indels 8; Gaps 3;
Qy 2 MALTEEKPIRPIANPPSIWGOFLIYQKQVGEQVQVNDLKKEVRQLLKALDIPMKH 61
Db 1 MAASSADKCRPLANFHSVNGYHFLSTHETNQEKVEDEYKETIRKMLVETCD---NS 57
Qy 62 ANLLKLIDEIQRLGIPYHFEREIDHALQCIYETTYGDNWNGDRS----SLWFLMRKQGY 117
Db 58 TQKLVLDAMQRLGVAYHFDNEIETSQNIPTDASSKQNDNDNNLYVWSLRFPLVRQGHY 117
Qy 118 VTCDVFNKYKNGAFKQSLANDVEGELLEZEATSMRVPGELIILEDALGFTSRSLINTK 177
Db 118 MSSDVFQKQTNQDGKFKETLNDVQGLLSLYEASHLRVNRNEEILEEALTFTTTHLESIV 177
Qy 178 DAFSTNPALFTEIQORALKQPIWKLPLRIEAAQYIPFYQQQDSHNKTLKLAKLEFNLLQ 237
Db 178 NLNNNNLSKVEGEALTOPIRMTLPRNGARKYISYENDDAHHLHLKPKALDPMLOK 237

Qy	238	LHKEELSHVCKWKWKAFDIKKNAPCLDRIVECYFWGLSGGYEPQYSRARVFFTKAVAVIT	297
Dd	238	FHQRELSDLTRWMKDOLFANKYPYADRILVECYFMIJGVFFPKYSRARKMWTKLNLTS	297
Qy	298	LIDDTYDAYGYYEELKIETEAVERWSIICLDLTPWVKPIYKLFWDVTYTEMEEFLEAKGR	357
Dd	298	IIDDTFAYATFDELUTFNDAIQORANNAIDSIQPMRPAYQALDITYSEMQVJSUKGB	357
Qy	358	TDLFNCCKGEFYKFVRNINLVAKWANE-CHIPETTEEHDPVWITIGANLLTTTCVLGMSD	416
Dd	358	LDRVYAKNMKKLVRAYPKETQWLNDCDHIKPKEEQOVENAIVSAGYMMISITTLVGIEB	417
Qy	417	IFTYESVEMWASAPLPFRYSGTGLRRLLDMTHKAEQRKHSSSLESVMKYNNVNEEVA	476
Dd	418	FISHETFELWNESVIVRASALIATAAMDIVGHEDQEQRGHVASLIECYMKDYOGASKQET	477
Qy	477	QTLTYKEVEDVWKODINREYLTKNTPRPLLMAVILCQFLVYQYAGKNFTRMGDYBKHL	536
Dd	478	YIKELKEVTNAWKODINKQFSRPTPEVPMFLERVLNLT RVADLTLYKEKDIYSTAKGKLKNM	537
Qy	537	IKSLLVYPMISI	547
Dd	538	INPILIESVKI	548

RESULT 4

```

US-09-903-012-52
; Sequence 52, Application US/09903012
; Patent No. US20020094557A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20020094557A11, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/903,012
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 52
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Lycopersicon esculentum
US-09-903-012-52

```

Query Match	38.6%;	Score 1123;	DB 9;	Length 548;
Best Local Similarity	38.8%;	Pred. No. 4e-95;		
Matches 214;	Conservative 123;	Mismatches 206;	Indels 8;	Gaps 3;
Qy	2	MALTEEKPIRPIANPPPSIMGDQLIYKQVQGVQVQVINDLKKEVRQLLKALDIPMKH	61	
Db	1	MAASSADKCRFLANFHPSPVGMVGHLSYHETITNQEKVDEVKETIRKQVETCD--NS	57	
Qy	62	ANLLKLDEIORLGIPIYHFEREIDHALOCITYETYGDNWNGDRS---SLWFRLMKQGY	117	
Db	58	TQKLVLDAMQELGVAYHFDNEITSIGNIFDASSKQNDNNLVYVSLRFLVYRQGHY	117	
Qy	118	VTCDVFNKYDKNGAFKOSLANDVEGLELEIYKATSRMVPVGLIILEDALGPTRSRLSIMTK	177	
Db	118	MSSDVFKQFTNQDGKFKELLTNDVQGLUSLYEASHLRVNRNEILIEEAUFTTTTHLESIVS	177	
Qy	178	DAFSTNPALFTEIORALKQPIKWLKRLPRIEAAQYTFPYQQQDSHNKTLTKLAKLEFNLLQS	237	

[illegible]

RESULT 5

```

US-09-900-797-52
; Sequence 52, Application US/09900797
; Publication No. US20030087406A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20030087406A1, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/900,797
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 52
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Lycopersicon esculentum
US-09-900-797-52

```

		Query Match	38.6%;	Score 1123;	DB 10;	Length 548;	
		Best Local Similarity	38.8%;	Pred. No. 4e-95;			
		Matches 214;	Conservative 123;	Mismatches 206;	Indels 8;	Gaps 3;	
Qy	2	MALTEEXPIRPIANFPFSIWGDOFLIYKQVQGVQEI VNDLKEEVQOLLKEALDIPMKH	61				
		: : : : : : : : :					
Db	1	MAASSADKCRLANFHSVWGTHFLSYTHITNQEKVEVDYEKETIRKMVLVETCD---NS	57				
		: : : : : : : : :					
Qy	62	ANLLKLIDETQRIGIPYHFEREIDHALOCTIYTYGDNNWGRS-----SLMFLRMKGOYY	117				
		: : : : : : : : :					
Db	58	TQKLVLDAMQRLGVAHYFDNEITSQNI F DASSKQNDDNMLYVVLSFRFLVRQQGHY	117				
		: : : : : : : : :					
Qy	118	VTCDFVNMYKDNGAFKQSILANDVEGLLELYEATSMVPGEIIILEDALGFTRSRLSIMTK	177				
		: : : : : : : : :					
Db	118	MSSDVFKQFTNQDGKFKEITLNDVOGLLSLYEASHLAVRNIEELILEALTFTTTHLESIVS	177				
		: : : : : : : : :					
Qy	178	DAFSTNPALFTTEIORALKQPILWKSLPRIEAAQYIPFYQQODSHNKTKLLKALEFNLLQS	237				
		: : : : : : : : :					
Db	178	NLSNNNSLKVEVEGALQTQIRMTLPBMGARKYISLYENNDAHHLHKFAKLFDMQLQK	237				
		: : : : : : : : :					

Qy 296 ITLIDDTVDAGTYBELKIFTEAVERWSITCLDTLPEYMKPIYKLPMDYTEMEEFLAKE 355
Db 305 ISIVDDTDFDAYGTVELEAYTDIAQRWDINEIDRLDPYMKISYKAILDLNDYDEKELS 354
Qy 356 GRDTLFCNGKEFVKFVRNLWEAKWANEHGHIPPTTEHDPVVIITGGANLTTTCYLGMS 415
Db 365 GRSDVVHVAKERKMEIVRNFYFEAKWFIEGYPPVSEYLSNALATYLLTTSYLGMK 424
Qy 416 DIFTKESVEWAVSAPPLFRYSIGILGRRLNDLTHKAEQERKSHSSSSLESYMKYNNVEEY 475
Db 425 SA-NQDFEWLAKNPKILEANVTLCRVIDDIAITYEVEKGRGQIATGIECYMRDYGSTK 483
Qy 476 AQTLLYKEVEDVKDINREYLTNNIPRLLMAVIVLCQFLEVQYA-GKDNFTRMGDEYK 534
Db 484 AMEKFOEMAETAKWQNEGILRPTPVSTBILTRILNLARIIDVTYKHODGYTHPEKVLK 543
Qy 535 HLIKSLLVYPMXI 547
Db 544 PHIIALLVDSIEI 556

RESULT 12
US-09-887-586A-2
; Sequence 2, Application US/09887586A
; Patent No. US20020094556A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20020094556A1, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/887,586A
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR FILING DATE: 1999-04-22
; PRIOR FILING DATE: 1999-04-22
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Nicotiana tabacum
US-09-887-586A-2

Query Match 35.1%; Score 1020.5; DB 9; Length 548;
Best Local Similarity 40.0%; Pred. No. 1.3e-85;
Matches 222; Conservative 102; Mismatches 202; Indels 29; Gaps 9;
Qy 6 BEKPIRPIANPPSPISWGDQFLIY--OKQVEQGVQIIVNDLKEVRQQLKEALDIPMKHAN 63
Db 10 EEEIYRPVADFSPLWGDQFLSFKINQVAEYAEIABKEQTRNML---LATGMKLAD 66
Qy 64 LLKLIDEIQRIGIPYHFEREIDHALQCIYETVYGDWNGD-----RSSLWFLMRKQGYV 118
Db 67 TLNLDITRIGISYHFEKEIDLDQIY-----NQNSCNDLCTSGALQFRLRQHGFI 121
Qy 119 TCDVFNNYKNGAFKQSLANDVEGLLELYEATSMRVPGEEIILEDALGFTSRSLSTMKD 178
Db 122 SPEIFSKFDENGKFKESLASDVLGLNLVEASHVTRHADDILEDALAFSTHLE----- 176
Qy 179 AFTSNPALFT----EIQRALKQPLWKLPRLEAAQYI-PFYQQODSHNKTLLKLALEFN 233
Db 177 --SAAPHLKSPLEQVTHALEQCLHKGVPVETRFPISSIIDYKESQKNVLLRFAKLDFN 234
Qy 234 LLQSLHKEELSHVCKWKAFFDIKNAPCLDRIVECYFWGLSGYEPQYSRARVFTTKAV 293
Db 235 LLQMLHKEELAQVSRWKKDLDFVTTLFYARDRVVECYFWALGVYFEPQYSQARVMLVKTI 294

Qy 294 AVITLIDDTVDAGTYBELKIFTEAVERWSITCLDTLPEYMKPIYKLPMDYTEMEEFLA 353
Db 295 SMISIVDDTDFDAYGTVELEAYTDIAQRWDINEIDRLDPYMKISYKAILDLNDYDEKELS 354
Qy 354 KEGRTDLFCNGKEFVKFVRNLWEAKWANEHGHIPPTTEHDPVVIITGGANLTTTCYLG 413
Db 355 SAGSHIVCHAIERKMEIVRNFYFEAKWFIEGYPPVSEYLSNALATYLLTTSYLG 414
Qy 414 MSIDFTKESVEWAVSAPPLFRYSIGILGRRLNDLTHKAEQERKSHSSSSLESYMKYNNVE 473
Db 415 MKSA-TEQDFEWLAKNPKILEASVTCRVIDDIAITYEVEKGRGQIATGIECCMRDYGIST 473
Qy 474 EYAQTLTYKEVEDVKDINREYLTNNIPRLLMAVIVLCQFLEVQYA-GKDNFTRMGDE 532
Db 474 KEAWAKFQMAETAKWQNEGILRPTPVSTBILTRILNLARIIDVTYKHODGYTHPEK 533
Qy 533 YKHLIKSLLVYPMXI 547
Db 534 LKPHIINLLVDSIKI 548

RESULT 13
US-09-903-012-2
; Sequence 2, Application US/09903012
; Patent No. US20020094557A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20020094557A1, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/903,012
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR FILING DATE: 1998-09-18
; PRIOR FILING DATE: 1999-04-22
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Nicotiana tabacum
US-09-903-012-2

Query Match 35.1%; Score 1020.5; DB 9; Length 548;
Best Local Similarity 40.0%; Pred. No. 1.3e-85;
Matches 222; Conservative 102; Mismatches 202; Indels 29; Gaps 9;
Qy 6 BEKPIRPIANPPSPISWGDQFLIY--OKQVEQGVQIIVNDLKEVRQQLKEALDIPMKHAN 63
Db 10 EEEIYRPVADFSPLWGDQFLSFKINQVAEYAEIABKEQTRNML---LATGMKLAD 66
Qy 64 LLKLIDEIQRIGIPYHFEREIDHALQCIYETVYGDWNGD-----RSSLWFLMRKQGYV 118
Db 67 TLNLDITRIGISYHFEKEIDLDQIY-----NQNSCNDLCTSGALQFRLRQHGFI 121
Qy 119 TCDVFNNYKNGAFKQSLANDVEGLLELYEATSMRVPGEEIILEDALGFTSRSLSTMKD 178
Db 122 SPEIFSKFDENGKFKESLASDVLGLNLVEASHVTRHADDILEDALAFSTHLE----- 176
Qy 179 AFTSNPALFT----EIQRALKQPLWKLPRLEAAQYI-PFYQQODSHNKTLLKLALEFN 233
Db 177 --SAAPHLKSPLEQVTHALEQCLHKGVPVETRFPISSIIDYKESQKNVLLRFAKLDFN 234
Qy 234 LLQSLHKEELSHVCKWKAFFDIKNAPCLDRIVECYFWGLSGYEPQYSRARVFTTKAV 293
Db 235 LLQMLHKEELAQVSRWKKDLDFVTTLFYARDRVVECYFWALGVYFEPQYSQARVMLVKTI 294

QY	294	AVITLIDDTYDAYGTYTELKIPTBAVERWSITCLDTLPEYMKPIYKLFMDTYTEMEEFIA	353
Db	295	SMISIVDDTFDAYGTVKELEAYTDAIQWDINEIDRLPDYMKISYKAILDLYKDEKELS	354
QY	354	KEGRDILFNCGKEFVKEVRNLMVEAKWANECHITTEEHDPVVIITGGANLITTCYLG	413
Db	355	SAGSHIVCHAIERNKEVVRNINVESTWFIQYTPVSEYLSNALATTTYVLTATSYLG	414
QY	414	MSDIFTKESVEWAVSAPPLFRYSGILGRRLNDLMTKAEQERKHSSSLESYMKYNNVNE	473
Db	415	MKSA-TEQDPFWLSKNPKILEASVILCRVIDDTATYEVEKSRGQIATGIECCMRDYGIST	473
QY	474	EYAOITLYKEVEDVWKDINREYLTITKNIPRPLMAVIVLCQFLEVQYA-GKDNFTRMGDE	532
Db	474	KEAMAKFQNMATAWKDINEGLLRPTVPVSTBELTILNLARIVEVTYIHNLDGYTHPEKV	533
QY	533	YKHLIKSLLVYPMSI	547
Db	534	LKPHIINLLVDSIKI	548

Search completed: June 3, 2004, 16:38:54
Job time : 52 secs

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: U23206; AAA93064.1; --
 DR PIR: S68365; S68365.
 DR HSP: Q40577; SEAU.
 DR InterPro: IPR008930; Terp_cyc_toroid.
 DR InterPro: IPR001906; Terp_synth-like.
 DR InterPro: IPR005630; Terpene synth C.
 DR InterPro: IPR008949; Terpenoid synth.
 DR Pfam: PF01397; Terpene synth; I.
 DR Pfam: PF03936; Terpene synth_C; 1.
 DR Lyase; Multigene family.
 FT ACT_SITE 451 BY SIMILARITY.
 FT ACT_SITE 527 BY SIMILARITY.
 FT ACT_SITE 531 BY SIMILARITY.
 SQ SEQUENCE 554 AA; 64137 MW; 59D6922DEDF9DCAF CRC64;

Query Match 39.3%; Score 1143; DB 1; Length 554;
 Best Local Similarity 42.1%; Pred. No. 2.3e-70;
 Matches 228; Conservative 106; Mismatches 197; Indels 10; Gaps 5;

QY 10 IRPIANPPSIWGWQFL-IYOKQVQGVQVWDLKKEVRQLLKEALDIPMKHA-NLLKL 67
 DB 21 MRKADQPSIWGDLFLNCPDNDAETKHKHQLEKVRKIV-APMANSTQKLA 76
 QY 68 IDEIQRIGIPYHFEREIDHALQCIYETYGDNWG-DRSSLWFLMRKQGYVTCDFVNNY 126
 DB 77 IDSQRLGVSHFTKEIEDELENIYHNNDAAENDLYTTSIRPFLLRHGYNSVCDVFNKF 136
 QY 127 KDNKAFQSLANDVEGLLELYEATSMRVPGEIILEDALGFTSRSLSIMTKDAFSPAL 196
 DB 137 KDEQGNFKSSVTSVGRGLLELYQASLYRVHGEDILDEAISFTTHLSLAVA-SLDHPL 193
 QY 187 FTEQRALKQPLWKLPRLEAAQYIPFYQQQDSHNKTLKLAKLFNLQSLHKEELSHV 246
 DB 194 SEEVSHALKQSIRGLPRVEARHYLSVQDIESHNKALLEFAKIDFNMLQFLHRELSEI 253
 QY 247 CKWKAFDIKKNAPCLDRIVECEYFVGLSGYEPQYSRVRVFTKAVAVITLDDTYDAY 306
 DB 254 CRWKKDLDFQKLPYARDRVVGVYFISGVYFEPQYSLGRKMLTKVIAMASIVDDTYDSY 313
 QY 307 GTYEELKIFTEAVERWSITCLDTLPYMKPIYKLFMDTYTEMEEFLLAKEGRDILFNGKE 366
 DB 314 ATYEELIPYTNARWDIKCIDELPEYMKPSYKALLDYVEEMVQLVAEHGQYRVYAKN 373
 QY 367 FVKEFVNLVBEAKWANEHGIPTTEBHDVPIITGGANLLTTTCYLGMSDIFTKESVEWA 426
 DB 374 AMRLAQSVLVEARWTLQNYKSPFEFKANALPTCGYAMLAITSFVGMGDIPTPEFKWA 433
 QY 427 VSAPFLFRYSIGILGRRLNDLMTKAEQERKHSLSSESYMKYNNVEEYAQTLIYKEVED 486
 DB 434 ASDPKIIQASTIICRFWDDVAEHKFKRRRDDCSAIECYMBEYGVTAQAYDVFNKHVES 493
 QY 487 VKKDINREYLTNNIPRPLMAVILYQFLEVOYAGKNFTMGDEYKHLIKSLIYVPM 546
 DB 494 AWKDLNQEFKPTTEMTPEVLNRLARVMDVLYREGDGYTYVYVGAAGKGTSLLEPIA 553
 QY 547 I 547
 DB 554 L 554

RESULT 3
 DCS2 GOSAR
 ID DCS2 GOSAR STANDARD; PRT; 554 AA.
 AC Q39761;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE (+)-delta-cadinene synthase isozyme XC14 (EC 4.2.3.13) (D-cadinene synthase).

DB 21 IRPKADQPSIWGWQFLNCPDNDAETKHKHQLEKVRKIV-APMANSTQKLA 76
 QY 68 IDEIQRIGIPYHFEREIDHALQCIYETYGDNWG-DRSSLWFLMRKQGYVTCDFVNNY 126
 DB 77 IDSQRLGVSHFTKEIEDELENIYHNNDAAENDLYTTSIRPFLLRHGYNSVCDVFNKF 136
 QY 127 KDNKAFQSLANDVEGLLELYEATSMRVPGEIILEDALGFTSRSLSIMTKDAFSPAL 196
 DB 137 KDEQGNFKSSVTSVGRGLLELYQASLYRVHGEDILDEAISFTTHLSLAVA-SLDHPL 193
 QY 187 FTEQRALKQPLWKLPRLEAAQYIPFYQQQDSHNKTLKLAKLFNLQSLHKEELSHV 246
 DB 194 SEEVSHALKQSIRGLPRVEARHYLSVQDIESHNKALLEFAKIDFNMLQFLHRELSEI 253
 QY 247 CKWKAFDIKKNAPCLDRIVECEYFVGLSGYEPQYSRVRVFTKAVAVITLDDTYDAY 306
 DB 254 CRWKKDLDFQKLPYARDRVVGVYFISGVYFEPQYSLGRKMLTKVIAMASIVDDTYDSY 313
 QY 307 GTYEELKIFTEAVERWSITCLDTLPYMKPIYKLFMDTYTEMEEFLLAKEGRDILFNGKE 366
 DB 314 ATYEELIPYTNARWDIKCIDELPEYMKPSYKALLDYVEEMVQLVAEHGQYRVYAKN 373
 QY 367 FVKEFVNLVBEAKWANEHGIPTTEBHDVPIITGGANLLTTTCYLGMSDIFTKESVEWA 426
 DB 374 AMRLAQSVLVEARWTLQNYKSPFEFKANALPTCGYAMLAITSFVGMGDIPTPEFKWA 433
 QY 427 VSAPFLFRYSIGILGRRLNDLMTKAEQERKHSLSSESYMKYNNVEEYAQTLIYKEVED 486
 DB 434 ANDPKIIQASTIICRFWDDVAEHKFKRRRDDCSAIECYMBEYGVSAQAYDVFNKHVES 493
 QY 487 VKKDINREYLTNNIPRPLMAVILYQFLEVOYAGKNFTMGDEYKHLIKSLIYVPM 546
 DB 494 AWKDLNQEFKPTTEMTPEVLNRLARVMDVLYREGDGYTYVYVGAAGKGTSLLEPIA 553
 QY 547 I 547
 DB 554 L 554

RESULT 2
 DCS1 GOSAR
 ID DCS1 GOSAR STANDARD; PRT; 554 AA.
 AC Q39761;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE (+)-delta-cadinene synthase isozyme XC1 (EC 4.2.3.13) (D-cadinene synthase).
 OS Gossypium arboreum (Tree cotton).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
 OX NCBI_TaxID=29729;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nanking.
 RX MEDLINE=96132653; PubMed=8554317;
 RA Chen X.-Y., Chen Y., Heinstein P., Davissson V.J.;
 RT "Cloning, expression, and characterization of (+)-delta-cadinene
 synthase: a catalyst for cotton phytoalexin biosynthesis."
 RL Arch. Biochem. Biophys. 324:255-266(1995).
 CC -!- FUNCTION: Responsible for the cyclization of trans,trans-farnesyl
 diphosphate (FPP) to (+)-delta cadinene.
 CC -!- CATALYTIC ACTIVITY: 2-trans,6-trans-farnesyl diphosphate = (+)-
 delta-cadinene + diphosphate.
 CC -!- PATHWAY: Phytoalexins gossypol and lacinilene C biosynthesis;
 CC first (committed) step.
 CC -!- SIMILARITY: Belongs to the terpene synthase family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way

QY 3 ALTEKPIRPIANFPBSIWGDOFLY-QKQVQGVQIIVNDLKKEVQRLKKEALDIPMKH 61
Db A I S S E N - R P K A D F P E G I W G D F I I C P D T D I A A T E L Q Y E E L K A Q V K R K I M E P V D - - - D S 69
QY 62 ANLLKLDI E I O R G L G I P Y F E R E I D H A L O C I Y E T Y G N W G D - - - - R S L M F R L M R K O G 115
Db 70 N O K L P F I D A V Q L G V S Y F E K I E B E L E N I Y R - - - D T N N D A D T D Y T T A L P R L L R E H G 126
QY 116 Y V T C D V F N N Y K D K N G A F K Q S L A N D V E G L L E Y A T S M R V P G E I I L E D A L G F T R S L S I M 175
Db 127 F D I S C A F N K F K D E A G N F K A S I T S D V Q G L L E Y A S Y M V H G E I D I D E A I S P T T A Q L T L A 186
QY 176 T K D A F S T P A L F T E T O R A L K Q P L M K L P R I E A A Q I P P Y Q Q O D S H N K T I L K L A K L E F N L L 235
Db 187 L P - - - T L H P U S E Q V G H A L K Q S I R G L P R V E A R N F I S I Y O D L E S H N K S L L Q A K I D F N L L 243
QY 236 Q S L H K E E L S H V K W K A F D I K N A P C L R D R I V E C Y F W G L G S G V E P O Y S R A R V F F T K A V A V 295
Db 244 Q U L L H K E L S E I C R W W K D I D F T R K L P A R D R V V E G Y F W I N G V F E P O Y S L G R K M L T K V I A M 303
QY 296 I T L I D D T Y D A Y G T Y B E L K I F T E A V E R W S I T C L D T L P E Y M K P I Y K U F M D T Y T T E M E E F L A K E 355
Db 304 A S I V D D T Y D S Y A T Y D E L I P Y T N A I E R W D I K M Q N L P N Y M K I S Y K A L L N V Y E E M E Q L L A N Q 363
QY 356 G R T D L F N G C K E F V K P V R N L M V E A K W A N E G H I P T T E E H D P V I I T G G A N L L T T C V L G M S 415
Db 364 G R Q R V E A K A M I R L Q A Y L E A K W T H Q N Y K P T F E E F R D N A L P T S G Y A M L A I T A F V G M G 423
QY 416 D I F T K E S V E A V S A P P L F R Y S G I L G R R L N D L M T H A E Q R K H S S S L E S Y M K E Y N N E E Y 475
Db 424 E V I T P E T K W A S D P K I I K A S T I I C R F W D D I A E H K F N H R E D D C S A I E C Y M K Q Y G T A G E 483
QY 476 A Q T I Y K E V D W K D I N E Y L T K N I P R P L M A V I Y L C O F L E V Q A G K D N F T R M G D E Y K H 535
Db 484 A Y N E F N K H I E S S W K D V N E E F L K P T E M P T V L C R S L N A R V M D V L Y R E G D G Y T H V G K A A G 543
QY 536 L I K S L I V P M S I 547
Db 544 G I T S L L I D P I Q I 555
RESULT 5
DCS1_GOSHI STANDARD; PRT; 554 AA.
AC P93665;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE (+)-delta-cadinene synthase (EC 4.2.3.13) (D-cadinene synthase).
GN CDN1.
OS Gossypium hirsutum (Upland cotton).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
OX NCBI_TaxID=3635;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=cv. Westburn M;
RA Davis E.M., Chen Y.-S., Essenberg M., Pierce M.L.;
RT "cdna sequence of a (+)-delta-cadinene synthase gene induced in
RL Gossypium hirsutum L. by bacterial infection."
RL (in) Plant Gene Register PGR98-040.
RN [2]
RN PARTIAL SEQUENCE AND CHARACTERIZATION.
RP MEDLINE=96351891; PubMed=8728715;
RX Davis E.M., Tsuji J., Davis G.D., Pierce M.L., Essenberg M.;
RA "Purification of (+)-delta-cadinene synthase, a sesquiterpene cyclase
RT from bacteria-inoculated cotton foliar tissue."
RL Phytochemistry 41:1047-1055(1996).
CC -!- FUNCTION: Responsible for the cyclization of trans,trans-farnesyl
CC diphosphate (FPP) to (+)-delta cadinene.
CC -!- CATALYTIC ACTIVITY: 2-trans,6-trans-farnesyl diphosphate = (+)-

CC delta-cadinene + diphosphate.
CC -!- PATHWAY: Phytoalexins gossypol and lacinilene C biosynthesis;
CC first (committed) step.
CC -!- INDUCTION: By bacterial infection.
CC -!- SIMILARITY: Belongs to the terpene synthase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U88318; AAC12784.1; -.
CC HSP; Q40577; SEAU.
DR InterPro; IPR008930; Terp_cyc_toroid.
DR InterPro; IPR001906; Terp_synth-like.
DR InterPro; IPR005630; Terpene_synth.C.
DR InterPro; IPR008949; Terpenoid_synth.
DR Pfam; PF01397; Terpene_synth; 1.
DR Pfam; PF03936; Terpene_synth_C; 1.
KW Lyase.
FT ACT_SITE 451 451 BY SIMILARITY.
FT ACT_SITE 527 527 BY SIMILARITY.
FT ACT_SITE 531 531 BY SIMILARITY.
SQ SEQUENCE 554 AA; 64019 MW; 8BCC78AD8C5B816 CRC64;

Query Match 38.7%; Score 1126; DB 1; Length 554;
Best Local Similarity 41.4%; Pred. No. 3,3e-69;
Matches 224; Conservative 111; Mismatches 196; Indels 10; Gaps 5;

QY 10 I R I A N P P S I W G D Q F L - I Y K Q V E Q G V Q I V N D L K K E V R O L L K E A L D I P M K H A N L - I K L 67
Db 21 M R P K A D F Q P S I W G D F L N C P D K N I D A E T Q K R H Q L K E E V R K M I V - - - A P M A N S T L K L A F 76
QY 68 I D E I Q S L G I P Y F E R E I D H A L O C I Y E T Y G N W G - D R S S L M F R L M R K O G Y V T C D V F N N Y 126
Db 77 I D S V Q G L G V S Y H F T K E I E D E L E N I Y H N N D A E N D L Y T T S R F R L L R E H G F H V C S D V F N K F 136
QY 127 K D K N G A F K Q S L A N D V E G L L E Y A T S M R V P G E I I L E D A L G F T R S L S I M T K D A F S T N P A L 186
Db 137 K D E Q G N F K S S V T S D V R G L L E Y A S Y L R V H G E I D I D E A I S F T S N H L S L A V A - - - S L D H P L 193
QY 187 F T E I Q A L K O P L W K R L P R I E A A Q I P P Y Q Q O D S H N K T I L K L A K L E F N L L Q S L H K E E L S H V 246
Db 194 S E R V S H A L K Q S I R R G L P R V E A R H V L S Y Q D I E S H N K V L L E F A K I D F N M V Q L L H R K E L S E I 253
QY 247 C K W K A F D I K N A P C L R D R I V E C Y F W G L G S G V E P O Y S R A R V F F T K A V A I L I D D T Y D A Y 306
Db 254 S R W W K D L D F Q R K L P Y A R D R V V E G Y F W I S G Y F F P Q Y S L G R K O L T K V I A M A S I V D D T Y D S Y 313
QY 307 G T Y E E L K I F T E A V E R W S I T C L D T L P E Y M K P I Y K U F M D T Y T T E M E E F L A K E G R T D L F N C K G E 366
Db 314 A T Y E E L I P T N A E R W D I K I D E L P E Y M K P S Y K A L L D V Y E M E Q L V A E H G Q Y R V E Y A K N 373
QY 367 F V K E F V R N L M V E A K W A N E G H I P T T E E H D P V I I T G G A N L L T T C V L G M S D I P T K E S V M A 426
Db 374 A M I R L A Q S V L V E A R W T L Q N Y K P S F E E P K A N A L P T C G Y A M L A I T S F V G M G D I V T P E T F K W A 433
QY 427 V S A P P L F R Y S G I L G R R L N D L M T H A E Q R K H S S S L E S Y M K E Y N N E E Y A Q T L I Y K E V D 486
Db 434 A N D P K I I Q A S T I I C R F W D D V T E H K F H R R E D D C S A I E C Y M E E Y G V T A Q E A Y D V F N K H V E S 493
QY 487 V W K D I N R E Y L T T K N I P R P L M A V I Y L C O F L E V Q A G K D N F T R M G D E Y K H L I K S L I V P M S 546
Db 494 A W K D V N Q G F L K P T E M P T V L C R S L N A R V M D V L Y R E G D G Y T H V G K A A G G I T S L I E P I A 553
QY 547 I 547
Db 554 L 554


```

10  E8EIVRVADPSPSLWGDQFUSFDINQVABKYAKEI8ALKEQTRNML-----LATQWKUAD 66
64  LLKLIDIEORLIGIPYHFEREDHALQCIYETYGDNWNGD-----RSSLWFLMRKQGYV 118
67  TLNLIDIERLIGISYHFEKEIDILQIY-----NQNSCNCDLCTSALQFLLRHQGFNI 121
119 TCDVFNNYKDNKGAFAKOSLANDVGLLELYEATSNRVPEGEIILEDALGFTSRLSIMTKD 178
122 SP8IFSXFODENGKFKESLASDVLGLLNLYEASHVRTHADDILEDALAFSTHLE----- 176
179 AFSINPALFT-----EIQRALKPQLKRLPRIEAAQYI--PFYQOQDSHNKTIILKLAKLSPFN 233
177 --GAAPHKLGPLREQVTHAUEQCQLHKGVPVETRFFISSIYDKBQSKNNVLLRFAKLDNF 234
234 LLOSLHKEELSHVCWKWKAFFDKKNAPCLDRIVECYFWGLSGVGEPOYSRARVFTKAV 293
235 LLQWLHKEELAQVSRWKNOLDFTVTLPLVARDRVVECYFWALGVPEPOYSQARVMLVKTI 294
294 AVTLIDTVDAYGTYBELKFTFAVERWSITCLDTPLEYNKPIYKLFMDTYTEMBEFLA 353
295 SM8ISIVDDTDAYTVKAELEYTDAIQ8WDINEIDRLPDYMKISYKAILDLVKDYKEKLS 354
354 KE8RTDLFCNGKEPVKPFVRNLMVEAKWANGHPTTEEHDPVVIITOGANLLTTTCVLG 413
355 SAGRSHVTHAIERKMEKVRNINVESTWFI8GYTPPVSEYI8NALATMTYYIYLATTSYL 414
414 MSDIFTKESVEWAVSAPPLPRYSGLIGRLRLDMLTHKAEQ8RKHSS8SLESYMKVEYNNE 473
415 MKSA--TQDPEWL8KNPKILEASVII8VIDDTATYVEVEK8GQIATGIECCMRDYGIST 473
474 EYAQTLIYKEVEDYWKDINREYLTQKIPRELLMAVIYLCOFLEQVYA--GKDNFTRMGDE 532
474 KEAMAKFQNM8AETAKKDINEGLLRPTPV8TEFLPILNARIVEYTHI8LDGYTHPEKV 533
533 YKHLIK8LLVYPMSI 547
534 LKPHIINLLVDSIKI 548

```

RESULT 7

AD	CASS_RICCO	STANDARD;	PRT;	601 AA.
IC	PS9287;			
DT	28-FEB-2003	(Rel. 41, Created)		
DT	28-FEB-2003	(Rel. 41, Last sequence update)		
DT	28-FEB-2003	(Rel. 41, Last annotation update)		
DE	Casbene synthase, chloroplast precursor	(EC 4.2.3.8).		
OS	Ricinus communis	(Castor bean).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;			
OC	eurosid I; Malpighiales; Euphorbiaceae; Acalyphoideae; Acalyphaeae;			
OC	Ricinus.			
ON	NCBI_TaxID=3988;			
OX	[1]			
RN	SEQUENCE FROM N.A.			
RP	MEDLINE=94359958; PubMed=8078910;			
RA	Mau C.J., West C.A.;			
RT	"Cloning of casbene synthase cDNA: evidence for conserved structural			
RT	features among terpenoid cyclases in plants."			
RL	Proc. Natl. Acad. Sci. U.S.A. 91:8497-8501(1994).			
RL	[2]			
RP	CHARACTERIZATION.			
RP	MEDLINE=97115639; PubMed=8954576;			
RA	Hill A.M., Cane D.E., Mau C.J., West C.A.;			
RT	"High level expression of Ricinus communis casbene synthase in			
RT	Escherichia coli and characterization of the recombinant enzyme.";			
RT	Arch. Biochem. Biophys. 336:283-289(1996).			
CC	-!- FUNCTION: Catalyzes the cyclization of geranylgeranyl diphosphate			
CC	to casbene, a diterpene phytoalexin with antibacterial and			
CC	antifungal activity.			
CC	-!- CATALYTIC ACTIVITY: Geranylgeranyl diphosphate = casbene +			
CC	diphosphate.			
CC	-!- SUBCELLULAR LOCATION: Caloroplast.			

170	VTDLNSTALGLRLRLHGVVSSDVKAFKQNGQSCSNIQTDEIRGVNLFRASLI	229
154	RVPGEIILEDALGFTSRSLSIMTKDAFSTNP--ALFTEIQRALKQPLWLKELPRIEAAQYI	211
230	AFFGEKIMDEAEIFFSTKYL----KEALQKIPVSSLSREIGDVLGYGWHYTLPLREARNYI	285
212	PFVQOQDSHN-----KTLMLKLAKLFPNLLQSLHKEEISHVCKWKAFFDKKNAPCLRD	264
286	QVT-GQDTENTKSYVKSCKLELAKLEFNIFQSLQKRELESVLVRWVKESGFFPEMTFC-RH	343
265	RIYBVCYFWGLSGVYEQYGRARVFFKAVAVITLDDTDVAYGTYBELKFTFEAVSRWSI	324
344	RHVEYVTLASCIAFEQHSQFRGLGFAKTHLITVLDDMDTFTGVDELEFLTATMKRWDP	403
325	TCLDTLPEYMKPIYKLFMDTYTTEM--EFLAKBQRTDLFCNGKEFVKFEFVRLNLMVEAKWAN	383
404	SSIDCLPEYMKGVYIAVDTVNEMAREASEAQGR-DTLTVAREAWAYIDSYMQEARNIA	462
384	EGHIPTEHDHPVUITGGANLLTTTCYLGMSDIFTKESVEWAVSAPPLFRYSGILGRRL	443
463	TGYLPSDFDEYKGVKSCGHRISALQPIITM-DIPFPDHILKEVDPP-----SKL	511
444	NDL-----MTHKAEQERKHSSSLESYMKKEY-VNNEEYQAOTLYIKEYEDVYKDI	491
512	NDLACAILRLRGTRCYKADRRAGEEASISCVMKONPGVSEEDALDHINAMISDVIKGL	571
492	NREYIATK-NIPRLPLMAVILYLCQFLEVQYAGKDNFRMGDEYKHILIKSLIV	542
572	NWELLKPDINVPISAKKHAFDARAHYGYKYRGDGYSVANVETKSLVTRTL	623
RESULT 9		
TSD2 ABIGR		
ID	TSD2 ABIGR	STANDARD; PRT; 627 AA.
AC	Q24474;	
DT	28-FEB-2003	(Rel. 41, Created)
DT	28-FEB-2003	(Rel. 41, Last sequence update)
DT	10-OCT-2003	(Rel. 42, Last annotation update)
DE	Myrcene synthase, chloroplast precursor [EC 4.2.3.15].	
GN	AG2.2.	
OS	Abies grandis (Grand fir).	
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
OC	Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Abies.	
OC	NCBI_TaxId=46611;	
RP	[1]	
RN	SEQUENCE FROM N.A., AND CHARACTERIZATION.	
RX	MEDLINE=97413772; PubMed=9268308;	
RT	Bohmann J., Steele C.L., Groteau R.;	
RT	"Monoterpene synthases from grand fir (Abies grandis). cDNA isolation,	
RT	characterization, and functional expression of myrcene synthase, (-)-	
RT	(4S)-limonene synthase, and (-)-(-1S,5S)-pinene synthase.";	
RL	J. Biol. Chem. 272:21784-21792(1997)	
CC	-I- FUNCTION: Involved in defensive oleoresin formation in conifers in	
CC	response to insect attack or other injury. Involved in monoterpene	
CC	(C10) olefins biosynthesis.	
CC	-I- CATALYTIC ACTIVITY: Geranyl diphosphate = myrcene + diphosphate.	
CC	-I- COFACTOR: Manganese and potassium.	
CC	-I- PATHWAY: Oleoresinosis.	
CC	-I- SUBCELLULAR LOCATION: Chloroplast.	
CC	-I- INDUCTION: By wounding.	
CC	-I- SIMILARITY: Belongs to the terpene synthase family.	
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/)	
CC	or send an email to license@isb-sib.ch	
DR	EMBL; U87908; AAB71084.1; -	
DR	HSSP; Q40577; 5EAS.	
DR	InterPro; IPR008930; Terp_cyc_toreid.	

```

RESULT 8
TSIDL_ABIGR STANDARD; PRT; 628 AA.
ID TSIDL_ABIGR ID TSIDL_ABIGR STANDARD; PRT; 628 AA.
AC 024475;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Pinene synthase, chloroplast precursor (EC 4.2.3.14) (Beta-geraniolene
DE synthase) ((-)-(1S,5S)-pinene synthase).
GN AG3.18.
OC Abies grandis (Grand fir).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Abies.
NCBI_TaxID=46611;
RN [1]
SEQUENCE FROM N.A., AND CHARACTERIZATION.
MEDLINE=97413772; PubMed=9268308;
RA Bohlmann J., Steele C.L., Croteau R.;
RT "Monoterpene synthases from grand fir (Abies grandis). cDNA isolation,
RT characterization, and functional expression of myrcene synthase, (-)-
RT (4S)-limonene synthase, and (-)-(1S,5S)-pinene synthase.";
RT J. Biol. Chem. 272:21784-21792(1997).
CC -!- FUNCTION: Involved in defensive oleoresin formation in conifers in
CC response to insect attack or other injury. Involved in monoterpene
CC (C10) olefins biosynthesis. A mixture of alpha-and beta-pinene is
CC produced by this enzyme.
CC -!- CATALYTIC ACTIVITY: Geranyl diphosphate = pinene + diphosphate.
CC -!- COFACTOR: Manganese and potassium.
CC -!- PATHWAY: Oleoresinosis.
CC -!- SUBCELLULAR LOCATION: Chloroplast.
CC -!- INDUCTION: By wounding.
CC -!- SIMILARITY: Belongs to the terpene synthase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U87909; AAB71085.1; -.
CC HSP; Q40577; SEAT.
CC InterPro; IPR008930; Terp_cyc_toroid.
CC InterPro; IPR001906; Terp_synth-like.
CC InterPro; IPR005630; Terpene_synth_C.
CC InterPro; IPR008949; Terpenoid_synth.
CC Pfam; PF01397; Terpene_synth_1.
CC Pfam; PF03936; Terpene_synth_C; 1.
CC Lyase; Manganese; Trans it peptide; Chloroplast.
CC TRANSIT 1 ? CHLOROPLAST (POTENTIAL).
CC CHAIN ? 628 PINENE SYNTHASE.
CC ACT_SITE 524 524 BY SIMILARITY.
CC ACT_SITE 601 601 BY SIMILARITY.
CC ACT_SITE 605 605 BY SIMILARITY.
CC SEQUENCE 628 AA; 71505 MW; 23DBE788F3C8072C CRC64;
CC
CC Query Match 21.1%; Score 614; DB 1; Length 628;
CC Best Local Similarity 30.2%; Pred. No. 2.6e-34;
CC Matches 179; Conservative 104; Mismatches 237; Indels 72; Gaps 17;
CC
CC QY 1 TWALTEKPRPIANPPSPWSWGQFL-----IYQKQVQ-----GVQIVNDLKEV 47
CC DB 54 TTVVDDGVRRMGDFSHNSLWDDVQLSPTAVEKSYLERAKLIGEVKNFNMSLED 113
CC
CC QY 48 RQLLKEALDIPMKHANLLKLIDRIQIGIPYHFERIDHALQCIYETYGDNWG----- 101
CC DB 114 GELMSPFLNDLIQR---LWIVDSLERLGIHRHFKDEIKSALDYVYSWGNGIGCGRESV 169
CC
CC QY 102 ----DRSSLWFLMRKQGIYVTCDFNNYKDKNGAFKQS-----LANDVEGLLELYEATSM 153
CC

```



```

RESULT 14
ID Y061_CAEEL STANDARD; PRT; 724 AA.
YO61_CAEEL
AC AC P3460;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical 84.7 kDa protein ZK1098.1 in chromosome III.
ZK1098.1.
GN Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OC NCBI_TaxID=6239;
OX [1]
RN SEQUENCE FROM N.A.
RP STRAIN=Bristol N2;
RC MEDLINE=94150718; PubMed=7906398;
RX Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA Latvalle P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
RA Sims M., Smaldon N., Smith A., Smith M., Sonnhammer E., Staden R.,
RA Sultston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Wohlman P.;
RA "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
CC -!- SIMILARITY: Contains 2 WW domains.
CC -!- SIMILARITY: TO S.POMBE SPAC4D7.13.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z22176; CAA80142.1; -.
DR PIR; F88557; F88557.
DR WormPep; ZK1098.1; CE03847.
DR InterPro; IPR002713; FF.
DR InterPro; IPR001202; WW_Reps_WWP.
DR Pfam; PF01846; FF; 5.
DR Pfam; PF00397; WW; 2.
DR SMART; SM00441; FF; 5.
DR SMART; SM00456; WW; 2.
DR PROSITE; PS01159; WW_DOMAIN_1; 2.
DR PROSITE; PS50020; WW_DOMAIN_2; 2.
DR KW Hypothetical protein; Repeat.
FT DOMAIN 78 111 WW 1.
FT FT DOMAIN 123 156 WW 2.
SQ SEQUENCE 724 AA; 84664 MW; BCC7AACAD5AA4C20 CRC64;

Query Match
Best Local Similarity 4.4%; Score 126.5; DB 1; Length 724;
Matches 110; Conservativity 90; Mismatches 171; Indels 169; Gaps 31;

QY 27 IYQVQVEQVGEQIWNLDKKEVROLLKEALDIPM--KXANLLKLIIDEIQIGIPYHFEREI 84
: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db 182 VQOKXAE-----SLDIDKAMKATLASMPNVPPLPSEKKEESVNDVELK-----KRQS 228

QY 85 DHALQCIYETGPD-----NWNNGRRSILW-----FLMRKQGYVYTCDVNNYKDKNG-- 131
: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
Db 229 ERFRELRDKYNDGKITTNCNWDQAVKWIQNDRFRILNKVS--EKKQLFNAWKVGQKGE 286

QY 132 -----AFKQSLANDVGLLELYEATSMVPGBEILEDALGFTSRSLSTMTKDAFSTNP 184

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; AE016967; AAP56564.1; -
DR EMBL; L38402; AAB40952.1; -
DR InterPro; IPR000722; RNA_pol_A.
DR InterPro; IPR007080; RNA_pol_Rpb1_1.
DR InterPro; IPR007066; RNA_pol_Rpb1_3.
DR InterPro; IPR007083; RNA_pol_Rpb1_4.
DR InterPro; IPR007081; RNA_pol_Rpb1_5.
DR InterPro; IPR006592; RNA_polA_N.
DR Pfam; PF04997; RNA_pol_Rpb1_1; 1.
DR Pfam; PF0623; RNA_pol_Rpb1_2; 1.
DR Pfam; PF04983; RNA_pol_Rpb1_3; 1.
DR Pfam; PF05000; RNA_pol_Rpb1_4; 1.
DR Pfam; PF04998; RNA_pol_Rpb1_5; 1.
DR SMART; SM00663; RPOA_N; 1.
KW Transferase; DNA-directed RNA polymerase; Transcription;
KW Complete proteome.
SQ SEQUENCE 1286 AA; 145058 MW; FB8C242C2D4AC693 CRC64;

Query Match 4.3%; Score 124; DB 1; Length 1286;
Best Local Similarity 20.7%; Pred. No. 1.4;
Matches 133; Conservative 89; Mismatches 194; Indels 226; Gaps 37;

QY 2 MALTTE-KEIRPIANPPSIWGQFLIYQKVEGGVQIYNDLKKEVRQLLKEALDIPMK 60
Db 445 MFNEDGGQIKPIA---PNIKIAEQIAK--SERIWDIVDKIKRPFVLNRA---PTL 496
QY 61 HANLLKLIDEIQRIGIPYHFERID-----HALQCIYETYGDNMNGDRSL----- 106
Db 497 H-----RLGIQAEPKIVDGAIRLHPL--VTTAFNADFDGDMVAVPISKEA 543
QY 107 -----WRLMRKQGYV---TCD-VFNYY-----KDKGAFKQSLANDVEGLLEL 147
Db 544 VAERAIMLASWHILGPDKGPVATPDQMVGLGNVYLTTEKRNEKGEGLIFFSDPDQVILA 603
QY 148 YEATSRVPEIILEDALGFTSRSLRMTKDAFTNPALFTEIQRALKQPLWKELPRIEA 207
Db 604 YEAKQVSIHALIGLS-----TKCLTKPKPAKQIGIVITVTKAIMNSI---MP--EE 649
QY 208 AQYIPFYQQDGHNKTL-----KLAKLEFNLLQSLHKEELSHVCK-WWKA 252
Db 650 MAYL-----NDGNLLELDESIVFAGEDPKQKLAKRP--LYKPFQKTKLSKIEILYKN 702
QY 253 FDIKNAPCLRDRIVECYFWGLSGYE-----PQYSRVRVFTKAVAVITLID 300
Db 703 FPLQK-VPOVLDKIKE-----FGFKYSTLSSTTISVFDIPRYDNKQEIYITKANEMIAKLK 756
QY 301 DTYDAVGYEELKIFTEAVERWS-----ITCLDTLPEYMKPIYKLFMDT----- 344
Db 757 HMYQK-GLTDDERYTKVRLMADVNDVSRDIKEITRPEYKENSIVVIADSGARGNIS 815
QY 345 -YTE---MEEFLLAKGRTDLFNCGEFVKFEFVRNLMVEAKWANEHGIPTTEHDPVVIIT 400
Db 816 NTFQFGMRGLMSKS-----YNYDQIKSQVIRDTI-----EVP1--KHS---FIE 856
QY 401 GGANLLTTCYL-----GMSDIFTKESVEWAVSAPPLFRYSGILGRINDLMTKAE 452
Db 857 G-----LTINEYFNSSYGARKGMTDIAMKTS-----KSGVMTRKLVD-----AA 895
QY 453 QERKHSSSL-----ESYMKYNNVEEY-----AQT 478
Db 896 QEVINDSCNTNKGIVVSTITNSLDGGVVTLSERIVTRYITIDPIYDEKTKELLVDADT 955
QY 479 LIYKEVEDWMDINREYLTNIPRLLMAVIY-----LCQ 514
Db 956 LITSEL-----AEKIAKANVTALIRSPIYCSQSTKGLCQ 989

Search completed: June 3, 2004, 16:39:28
Job time : 24 secs